GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 15, 2003, 06:15:49 ; Search time 3999 Seconds
(without alignments)
10916.281 Million cell updates/sec

Title: US-09-043-944-5
Perfect score: 1500
Sequence: 1 gtttaattacccaagtttga.....taaaaaaaaaaaaaaaaaa 1500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1454.6	97.0	1461	3	CEU35660	U35660 Caenorhabdi
2	1311.8	87.5	1335	3	AF171064	AF171064 Caenorhab
3	631	42.1	27102	3	U41540	U41540 Caenorhabdi
4	276.2	18.4	1896	4	AF038936	AF038936 Bos tauru
5	268.2	17.9	2428	5	DRE132931	AJ132931 Danio rer
6	264.2	17.6	2607	5	D84427	D84427 Xenopus lae
7	261.6	17.4	1836	5	AY043492	AY043492 Gallus ga
8	261	17.4	1430	9	MMPRSNLNI	Z71333 M.murinus p
9	258.4	17.2	1407	10	RATS182	D82578 Rattus norv
10	255.2	17.0	1407	10	D82363	D82363 Rattus norv
11	252.2	16.8	1241	5	CCPS1	Y17128 Cyprinus ca
12	248.2	16.5	1392	9	HSU40379	U40379 Human prese
13	248.2	16.5	1404	6	AX195267	AX195267 Sequence
14	248.2	16.5	1488	6	AR086330	AR086330 Sequence
15	248.2	16.5	2764	6	AR095621	AR095621 Sequence
16	248.2	16.5	2765	6	AX244146	AX244146 Sequence
17	248.2	16.5	2765	6	AX244648	AX244648 Sequence
18	248.2	16.5	2765	9	HUMS182R	L42110 Homo sapien
19	248.2	16.5	2765	11	G27112	G27112 human STS S
20	248.2	16.5	2791	6	AR060154	AR060154 Sequence
21	248.2	16.5	2791	6	AR087309	AR087309 Sequence
22	248.2	16.5	2791	6	AR134496	AR134496 Sequence
23	248.2	16.5	2792	6	AR144597	AR144597 Sequence
24	248.2	16.5	2811	9	BC011729	BC011729 Homo sapi
25	247	16.5	1404	6	AX418176	AX418176 Sequence
26	246.6	16.4	1914	6	AR082573	AR082573 Sequence
27	246.6	16.4	2791	6	AR060024	AR060024 Sequence
28	246.6	16.4	2791	6	AR087179	AR087179 Sequence
29	246.6	16.4	2791	6	AR134366	AR134366 Sequence
30	246.6	16.4	2791	6	AR144467	AR144467 Sequence
31	246.6	16.4	3056	9	HUMPS1MRNA	L76517 Homo sapien
32	246.6	16.4	3087	6	AR060026	AR060026 Sequence
33	246.6	16.4	3087	6	AR087181	AR087181 Sequence
34	246.6	16.4	3087	6	AR134368	AR134368 Sequence
35	246.6	16.4	3087	6	AR144469	AR144469 Sequence
36	243.2	16.2	1962	6	AR060155	AR060155 Sequence
37	243.2	16.2	1962	6	AR087310	AR087310 Sequence
38	243.2	16.2	1962	6	AR134497	AR134497 Sequence
39	243.2	16.2	1964	6	AR144598	AR144598 Sequence
40	243.2	16.2	1964	10	MUSS1PR	L42177 Mus musculu
41	243.2	16.2	2016	10	BC030409	BC030409 Mus muscu
42	243.2	16.2	2681	6	A63557	A63557 Sequence 10
43	240.8	16.1	1410	10	AF149111	AF149111 Mus muscu
44	239	15.9	1484	9	HSU40380	U40380 Human prese
45	238.2	15.9	1430	5	AY043493	AY043493 Gallus ga
46	237.4	15.8	1254	9	AF416717	AF416717 Homo sapi
47	233.6	15.6	1929	6	AR060025	AR060025 Sequence
48	233.6	15.6	1929	6	AR087180	AR087180 Sequence
49	233.6	15.6	1929	6	AR134367	AR134367 Sequence
50	233.6	15.6	1929	6	AR144468	AR144468 Sequence
51	228.2	15.2	2490	5	D84428	D84428 Xenopus lae
52	226.2	15.1	1764	3	DMU78084	U78084 Drosophila
53	226.2	15.1	1809	3	DMU77934	U77934 Drosophila
54	226.2	15.1	1894	3	AF017024	AF017024 Drosophil
55	226.2	15.1	1895	6	AR060183	AR060183 Sequence
56	226.2	15.1	1895	6	AR087338	AR087338 Sequence
57	226.2	15.1	1895	6	AR134525	AR134525 Sequence
58	226.2	15.1	2072	3	AY061316	AY061316 Drosophil
59	225.8	15.1	1404	6	AX418177	AX418177 Sequence
60	225.8	15.1	1705	9	HSU34349	U34349 Human seven
61	225.8	15.1	2276	6	AR086354	AR086354 Sequence
62	224.6	15.0	1404	6	AX418178	AX418178 Sequence
63	224.2	14.9	1346	6	AX195269	AX195269 Sequence
64	224.2	14.9	1347	6	AX418183	AX418183 Sequence
65	224.2	14.9	1983	6	E58793	E58793 Method for

66	224.2	14.9	2144	6	E58791	E58791 Method for
67	224.2	14.9	2229	9	HUME51P	L44577 Homo sapien
68	224.2	14.9	2236	6	AX244148	AX244148 Sequence
69	224.2	14.9	2236	6	AX244650	AX244650 Sequence
70	224.2	14.9	2236	9	HUMSTM2R	L43964 Homo sapien
71	224.2	14.9	2285	6	AR060156	AR060156 Sequence
72	224.2	14.9	2285	6	AR087311	AR087311 Sequence
73	224.2	14.9	2285	6	AR134498	AR134498 Sequence
74	224.2	14.9	2296	6	AR144599	AR144599 Sequence
75	224.2	14.9	2517	9	BC006365	BC006365 Homo sapi

ALIGNMENTS

```
RESULT 1
CEU35660
LOCUS       CEU35660                1461 bp    mRNA    linear   INV 23-JAN-1996
DEFINITION  Caenorhabditis elegans membrane protein (sel-12) mRNA, complete
            cds.
ACCESSION   U35660
VERSION     U35660.1  GI:1109793
KEYWORDS    .
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 1461)
  AUTHORS   Levitan,D. and Greenwald,I.
  TITLE     Facilitation of lin-12-mediated signalling by sel-12, a
            Caenorhabditis elegans S182 Alzheimer's disease gene
  JOURNAL   Nature 377 (6547), 351-354 (1995)
  MEDLINE   96032531
   PUBMED   7566091
REFERENCE   2  (bases 1 to 1461)
  AUTHORS   Levitan,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-1995) Diane Levitan, Biochemistry, Columbia
            University, 630 W. 168th St., New York, NY 10032, USA
COMMENT     On Dec 9, 1995 this sequence version replaced gi:1000566.
FEATURES             Location/Qualifiers
     source          1. .1461
                     /organism="Caenorhabditis elegans"
                     /strain="Bristol var N2"
                     /db_xref="taxon:6239"
                     /map="LGX"
     gene            1. .1461
                     /gene="sel-12"
     CDS             1. .1386
                     /gene="sel-12"
                     /note="multiple membrane spanning domains"
                     /codon_start=1
                     /product="membrane protein"
                     /protein_id="AAA85511.1"
                     /db_xref="GI:1109794"
                     /translation="MPSTRRQQEGGGADAETHTVYGTNLITNRNSQEDENVVEEAELK
                     YGASHVIHLFVPVSLCMALVVFTMNTITFYSQNNGRHLLYTPFVRETDSIVEKGLMSL
                     GNALVMLCVVVLMTVLLIVFYKYKFYKLIHGWLIVSSFLLLFLFTTIYVQEVLKSFDV
                     SPSALLVLFGLGNYGVLGMMCIHWKGPLRLQQFYLITMSALMALVFIKYLPEWTVWFV
                     LFVISVWDLVAVLTPKGPLRYLVETAQERNEPIFPALIYSSGVIYPYVLVTAVENTTD
                     PREPTSSDSNTSTAFPGEASCSSETPKRPKVKRIPQKVQIESNTTASTTQNSGVRVER
                     ELAAERPTVQDANFHRHEEEERGVKLGLGDFIFYSVLLGKASSYFDWNTTIACYVAIL
                     IGLCFTLVLLAVFKRALPALQFPPFSPDSFFTFVPAGSSPHLLHKSLKSVYYINSLFLP
                     FLCIINFSIIS"
     polyA_site      1461
                     /gene="sel-12"
                     /note="20 A residues"
BASE COUNT      379 a    309 c    310 g    463 t
ORIGIN
```

Query Match 97.0%; Score 1454.6; DB 3; Length 1461;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1457; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      22 ATGCCTTCCACAAGGAGACAACAGGAGGGCGGAGGTGCAGATGCGGAAACACATACCGTT 81
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGCCTTCCACAAGGAGACAACAGGAGGGCGGAGGTGCAGATGCGGAAACACATACCGTT 60

Qy      82 TACGGTACAAATCTGATAACAAATCGGAATAGCCAAGAAGACGAAAATGTTGTGGAAGAA 141
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TACGGTACAAATCTGATAACAAATCGGAATAGCCAAGAAGACGAAAATGTTGTGGAAGAA 120

Qy     142 GCGGAGCTGAAATACGGAGCATCTCACGTTATTCATCTATTTGTGCCGGTGTCACTATGC 201
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCGGAGCTGAAATACGGAGCATCTCACGTTATTCATCTATTTGTGCCGGTGTCACTATGC 180

Qy     202 ATGGCTCTGGTTGTTTTTACGATGAACACGATTACGTTTTATAGTCAAAACAATGGAAGG 261
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGGCTCTGGTTGTTTTTACGATGAACACGATTACGTTTTATAGTCAAAACAATGGAAGG 240

Qy     262 CATTTACTATCACATCCTTTTGTCCGGGAAACAGACAGTATCGTTGAGAAGGGATTGATG 321
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CATTTACTATACACTCCTTTTGTCCGGGAAACAGACAGTATCGTTGAGAAGGGATTGATG 300

Qy     322 TCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCGTTCTGATGACAGTTCTGCTGATT 381
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCGTTCTGATGACAGTTCTGCTGATT 360

Qy     382 GTTTTCTATAAATACAAGTTTTATAAGCTTATTCATGGATGGCTTATTGTCAGCAGTTTT 441
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTTTTCTATAAATACAAGTTTTATAAGCTTATTCATGGATGGCTTATTGTCAGCAGTTTT 420

Qy     442 CTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAGAAGTTCTGAAAAGTTTCGATGTG 501
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAGAAGTTCTGAAAAGTTTCGATGTG 480

Qy     502 TCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTAACTATGGAGTTCTCGGAATGATG 561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTAACTATGGAGTTCTCGGAATGATG 540

Qy     562 TGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCA 621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCA 600

Qy     622 CTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAATGGACTGTGTGGTTTGTGCTGTTT 681
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAATGGACTGTGTGGTTTGTGCTGTTT 660

Qy     682 GTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACACCAAAAGGACCATTGAGATATTTG 741
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACACCAAAAGGACCATTGAGATATTTG 720

Qy     742 GTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCCCGGCGCTGATTTATTCGTCTGGA 801
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCCCGGCGCTGATTTATTCGTCTGGA 780

Qy     802 GTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAAACACGACAGACCCCCGTGAACCG 861
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAAACACGACAGACCCCCGTGAACCG 840

Qy     862 ACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTGGAGAGGCGAGTTGTTCATCTGAA 921
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTGGAGAGGCGAGTTGTTCATCTGAA 900

Qy     922 ACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAAAAGTGCAAATCGAATCGAATACT 981
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAAAAGTGCAAATCGAATCGAATACT 960

Qy     982 ACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGGAACGGGAGCTAGCTGCTGAGAGA 1041
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGGAACGGGAGCTAGCTGCTGAGAGA 1020

Qy    1042 CCAACTGTACAAGACGCCAATTTTCACAGGCACGAAGAGGAAGAGAGAGGTGTGAAACTT 1101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CCAACTGTACAAGACGCCAATTTTCACAGGCACGAAGAGGAAGAGAGAGGTGTGAAACTT 1080
```

```
Qy       1102 GGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCGGCAAGGCTTCATCGTACTTTGAC 1161
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCGGCAAGGCTTCATCGTACTTTGAC 1140

Qy       1162 TGGAACACGACTATCGCTTGTTATGTGGCCATTCTTATCGGTCTCTGCTTCACTCTTGTC 1221
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGGAACACGACTATCGCTTGTTATGTGGCCATTCTTATCGGTCTCTGCTTCACTCTTGTC 1200

Qy       1222 CTGCTCGCCGTCTTCAAACGAGCACTCCCGGCTCTGCAATTTCCATTTTCTCCGGACTCA 1281
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGCTCGCCGTCTTCAAACGAGCACTCCCGGCTCTGCAATTTCCATTTTCTCCGGACTCA 1260

Qy       1282 TTTTTTACTTTTGTACCCGCTGGATCATCACCCCATTTGTTACACAAGTCTCTCAAAAGT 1341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTTTTTACTTTTGTACCCGCTGGATCATCACCCCATTTGTTACACAAGTCTCTCAAAAGT 1320

Qy       1342 GTTTATTATATTAATTCTCTGTTTTTGCCATTTCTTTGCATCATCAACTTTTCGATTATA 1401
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTTTATTATATTAATTCTCTGTTTTTGCCATTTCTTTGCATCATCAACTTTTCGATTATA 1380

Qy       1402 TCTTGAGCGATCTCAAAGCTTTATTTTACATACCTATTTATTTTTGAACTTTGTCATTTA 1461
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTTGAGCGATCTCAAAGCTTTATTTTACATACCTATTTATTTTTGAACTTTGTCATTTA 1440

Qy       1462 AGTTATATAAATAATTTATTA 1482
             |||||||||||||||||||||
Db       1441 AGTTATATAAATAATTTATTA 1461
```

```
RESULT 2
AF171064
LOCUS       AF171064                1335 bp    mRNA    linear   INV 23-AUG-1999
DEFINITION  Caenorhabditis elegans presenilin SEL-12 (sel-12) mRNA, complete
            cds.
ACCESSION   AF171064
VERSION     AF171064.1  GI:5759132
KEYWORDS    .
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 1335)
  AUTHORS   Eimer,S. and Baumeister,R.
  TITLE     Caenorhabditis elegans presenilin SEL-12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1335)
  AUTHORS   Eimer,S. and Baumeister,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUL-1999) Genzentrum, LMU Munich, Feodor-Lynen-Str.
            25, Munich 81377, Germany
FEATURES             Location/Qualifiers
     source          1. .1335
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /chromosome="X"
     gene            1. .1335
                     /gene="sel-12"
     CDS             1. .1335
                     /gene="sel-12"
                     /function="facilitator of Notch receptors signaling"
                     /note="membrane protein; similar to Homo sapiens PS1 and
                     PS2; similar to SEL12 deposited in GenBank Accession
                     Number AAA85511"
                     /codon_start=1
                     /product="presenilin SEL-12"
                     /protein_id="AAD50991.1"
                     /db_xref="GI:5759133"
                     /translation="MPSTRRQQEGGGADAETHTVYGTNLITNRNSQEDENVVEEAELK
                     YGASHVIHLFVPVSLCMALVVFTMNTITFYSQNNGRHLLYTPFVRETDSIVEKGLMSL
                     GNALVMLCVVVLMTVLLIVFYKYKFYKLIHGWLIVSSFLLLFLFTTIYVQEVLKSFDV
                     SPSALLVLFGLGNYGVLGMMCIHWKGPLRLQQFYLITMSALMALVFIKYLPEWTVWFV
                     LFVISVWDLVAVLTPKGPLRYLVETAQERNEPIFPALIYSSGVIYPYVLVTAVENTTD
                     PREPTSSDSNTSTAFPGEASCSSETPKRPKVKRIPQKVQIESNTTASTTQNSGVRVER
                     ELAAERPTVQDANFHRHEEEERGVKLGLGDFIFYSVLLGKASSYFDWNTTIACYVAIL
                     IGLCFTLVLLAVFKRALPALPISIFSGLIFYFCTRWIITPFVTQVSQKCLLY"
BASE COUNT      345 a    291 c    301 g    398 t
ORIGIN
```

```
  Query Match             87.5%;  Score 1311.8;  DB 3;  Length 1335;
  Best Local Similarity   99.4%;  Pred. No. 1.9e-289;
  Matches 1327;  Conservative    0;  Mismatches    7;  Indels    1;  Gaps    1;

Qy         22 ATGCCTTCCACAAGGAGACAACAGGAGGGCGGAGGTGCAGATGCGGAAACACATACCGTT 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTTCCACAAGGAGACAACAGGAGGGCGGAGGTGCAGATGCGGAAACACATACCGTT 60

Qy         82 TACGGTACAAATCTGATAACAAATCGGAATAGCCAAGAAGACGAAAATGTTGTGGAAGAA 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACGGTACAAATCTGATAACAAATCGGAATAGCCAAGAAGACGAAAATGTTGTGGAAGAA 120

Qy        142 GCGGAGCTGAAATACGGAGCATCTCACGTTATTCATCTATTTGTGCCGGTGTCACTATGC 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGAGCTGAAATACGGAGCATCTCACGTTATTCATCTATTTGTGCCGGTGTCACTATGC 180

Qy        202 ATGGCTCTGGTTGTTTTTACGATGAACACGATTACGTTTTATAGTCAAAACAATGGAAGG 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGCTCTGGTTGTTTTTACGATGAACACGATTACGTTTTATAGTCAAAACAATGGAAGG 240

Qy        262 CATTTACTATCACATCCTTTTGTCCGGGAAACAGACAGTATCGTTGAGAAGGGATTGATG 321
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATTTACTATACACTCCTTTTGTCCGGGAAACAGACAGTATCGTTGAGAAGGGATTGATG 300

Qy        322 TCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCGTTCTGATGACAGTTCTGCTGATT 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCGTTCTGATGACAGTTCTGCTGATT 360

Qy        382 GTTTTCTATAAATACAAGTTTTATAAGCTTATTCATGGATGGCTTATTGTCAGCAGTTTT 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTTTCTATAAATACAAGTTTTATAAGCTTATTCATGGATGGCTTATTGTCAGCAGTTTT 420

Qy        442 CTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAGAAGTTCTGAAAAGTTTCGATGTG 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAGAAGTTCTGAAAAGTTTCGATGTG 480

Qy        502 TCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTAACTATGGAGTTCTCGGAATGATG 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTAACTATGGAGTTCTCGGAATGATG 540

Qy        562 TGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCA 600

Qy        622 CTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAATGGACTGTGTGGTTTGTGCTGTTT 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAATGGACTGTGTGGTTTGTGCTGTTT 660

Qy        682 GTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACACCAAAAGGACCATTGAGATATTTG 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACACCAAAAGGACCATTGAGATATTTG 720

Qy        742 GTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCCCGGCGCTGATTTATTCGTCTGGA 801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCCCGGCGCTGATTTATTCGTCTGGA 780

Qy        802 GTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAAACACGACAGACCCCCGTGAACCG 861
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        781 GTCATCTATCCCTACGTTCTTGTTACTGCGGTTGAAAACACGACAGACCCCCGTGAACCG 840

Qy        862 ACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTGGAGAGGCGAGTTGTTCATCTGAA 921
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        841 ACGTCGTCAGACTCAAATACTTCTACAGCGTTTCCTGGAGAGGCGAGTTGTTCATCTGAA 900

Qy        922 ACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAAAAGTGCAAATCGAATCGAATACT 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAAAAGTGCAAATCGAATCGAATACT 960
```

```
Qy        982 ACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGGAACGGGAGCTAGCTGCTGAGAGA 1041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGGAACGGGAGCTAGCTGCTGAGAGA 1020

Qy       1042 CCAACTGTACAAGACGCCAATTTTCACAGGCACGAAGAGGAAGAGAGAGGTGTGAAACTT 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCAACTGTACAAGACGCCAATTTTCACAGGCACGAAGAGGAAGAGAGAGGTGTGAAACTT 1080

Qy       1102 GGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCGGCAAGGCTTCATCGTACTTTGAC 1161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCGGCAAGGCTTCATCGTACTTTGAC 1140

Qy       1162 TGGAACACGACTATCGCTTGTTATGTGGCCATTCTTATCGGTCTCTGCTTCACTCTTGTC 1221
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGGAACACGACTATCGCCTGTTATGTGGCCATTCTTATCGGTCTCTGCTTCACTCTTGTC 1200

Qy       1222 CTGCTCGCCGTCTTCAAACGAGCACTCCCGGCTCTG-CAATTTCCATTTTCTCCGGACTC 1280
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1201 CTGCTCGCCGTCTTCAAACGAGCACTCCCGGCTCTGCCAATTTCCATTTTCTCCGGACTC 1260

Qy       1281 ATTTTTTACTTTTGTACCCGCTGGATCATCACCCCATTTGTTACACAAGTCTCTCAAAAG 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATTTTTTACTTTTGTACCCGCTGGATCATCACCCCATTTGTTACACAAGTCTCTCAAAAG 1320

Qy       1341 TGTTTATTATATTAA 1355
              |||||||||||||||
Db       1321 TGTTTATTATATTAA 1335
```

RESULT 3
U41540

```
LOCUS       U41540                 27102 bp    DNA     linear   INV 18-JUN-2002
DEFINITION  Caenorhabditis elegans cosmid F35H12, complete sequence.
ACCESSION   U41540
VERSION     U41540.1  GI:1109862
KEYWORDS    HTG.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 27102)
  AUTHORS   Waterston,R.
  TITLE     Genome sequence of the nematode C. elegans: a platform for
            investigating biology. The C. elegans Sequencing Consortium
  JOURNAL   Science 282 (5396), 2012-2018 (1998)
  MEDLINE   99069613
   PUBMED   9851916
REFERENCE   2  (bases 1 to 27102)
  AUTHORS   Nelson,J. and Gattung,S.
  TITLE     The sequence of C. elegans cosmid F35H12
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 27102)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-APR-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   4  (bases 1 to 27102)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2001) Department of Genetics, Washington
            University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
            Louis, MO 63110, USA
REFERENCE   5  (bases 1 to 27102)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2002) Department of Genetics, Washington
            University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
            Louis, MO 63110, USA
REFERENCE   6  (bases 1 to 27102)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-2002) Department of Genetics, Washington
            University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
            Louis, MO 63110, USA
COMMENT     Submitted by:
                  Genome Sequencing Center
                  Department of Genetics, Washington University
                  St. Louis , MO  63110, USA, and
                  Sanger Centre, Hinxton Hall
                  Cambridge CB10 IRQ, England
                  email: rw@nematode.wustl.edu and jes@sanger.ac.uk

            NOTICE:  This sequence may not be the entire insert of this clone.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we provide a small overlap between
            neighboring submissions.

            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate chemistry
            or covered by high quality data (i.e., phred quality >= 30); an
            attempt was made to resolve all sequencing problems, such as
            compressions and repeats; all regions were covered by sequence from
            more than one m13 subclone.

            For a graphical representation of this cosmid sequence and its
            analysis see:
            http://www.wormbase.org/db/seq/sequence?name=F35H12;class=Sequence

                        NEIGHBORING COSMID INFORMATION

            The 5' cosmid is ZC13, 200 bp overlap; the 3' cosmid is F53H8, 200
            bp overlap. Actual start of this cosmid is at base position 1 of
            F35H12; actual end is at 12696 of F53H8.

            NOTES:

            Coding seqences below are the result of integration and manual
            review of the following data : computer analysis using the program
            Genefinder (P. Green and L. Hillier, personal communication), the
            large scale EST projects of Yuji Kohara
            (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C.
            elegans ORFeome cloning project (http://worfdb.dfci.harvard.edu/),
            similarity to other proteins from BlastX analyses
            (http://blast.wustl.edu/), sequence conservation with C. briggsae
            using Jim Kent's WABA alignment program (Genome Research
            10:1115-1125, 2000), individual C. elegans GenBank submissions,
            and personal communications with C. elegans researchers. tRNAs
            are predicted using the program tRNAscan-SE (Lowe, T.M. and
            Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).
FEATURES                Location/Qualifiers
     source             1. .27102
                        /organism="Caenorhabditis elegans"
                        /strain="Bristol N2"
                        /db_xref="taxon:6239"
                        /chromosome="X"
                        /clone="F35H12"
     gene               2463. .3251
                        /gene="F35H12.6"
                        /note="for a graphical representation of this gene see:
                        http://www.wormbase.org/db/seq/sequence?name=F35H12.6;clas
                        s=Sequence"
     CDS                join(2463. .2552,3063. .3251)
                        /gene="F35H12.6"
                        /note="coded for by the following C. elegans cDNAs:
                        yk1092f08.3"
                        /codon_start=1
                        /product="Hypothetical protein F35H12.6"
                        /protein_id="AAM54184.1"
                        /db_xref="GI:21450554"
                        /translation="MESYQSGQYSESDSYSHEYSLNTVQNGLAGVPMWAIWSFVICAV
                        LIIAVLLLDYCVIRRNALGNTCCTKARQKRGTINAAHQKRSTNMLLNV"
     gene               5010. .7295
                        /gene="sel-12"
                        /note="for a graphical representation of this gene see:
                        http://www.wormbase.org/db/seq/sequence?name=F35H12.3;clas
```

```
                s=Sequence"
CDS             join(5010. .5063,5134. .5324,5629. .5838,5884. .6203,
                6265. .6348,6398. .6604,7027. .7295)
                /gene="sel-12"
                /note="C. elegans membrane protein (SEL-12) (SP:P52166);
                contains similarity to Pfam domain PF01080 (Presenilin),
                Score=918.2, E-value=7.4e-273, N=1; coded for by the
                following C. elegans cDNAs: yk216e1.3, yk221d3.3,
                yk221d3.5, yk231a7.3, yk231a7.5, yk400e8.3, yk400e8.5,
                yk452b9.3, yk452b9.5, yk499e3.3, yk573h4.3, yk600e12.3,
                yk674e3.3, AF171064, U35660"
                /codon_start=1
                /product="C. elegans SEL-12 protein (corresponding
                sequence F35H12.3)"
                /protein_id="AAK39230.1"
                /db_xref="GI:13775416"
                /translation="MPSTRRQQEGGGADAETHTVYGTNLITNRNSQEDENVVEEAELK
                YGASHVIHLFVPVSLCMALVVFTMNTITFYSQNNGRHLLYTPFVRETDSIVEKGLMSL
                GNALVMLCVVVLMTVLLIVFYKYKFYKLIHGWLIVSSFLLLFLFTTIYVQEVLKSFDV
                SPSALLVLFGLGNYGVLGMMCIHWKGPLRLQQFYLITMSALMALVFIKYLPEWTVWFV
                LFVISVWDLVAVLTPKGPLRYLVETAQERNEPIFPALIYSSGVIYPYVLVTAVENTTD
                PREPTSSDSNTSTAFPGEASCSSETPKRPKVKRIPQKVQIESNTTASTTQNSGVRVER
                ELAAERPTVQDANFHRHEEEERGVKLGLGDFIFYSVLLGKASSYFDWNTTIACYVAIL
                IGLCFTLVLLAVFKRALPALPISIFSGLIFYFCTRWIITPFVTQVSQKCLLY"
misc_feature    7532. .7533
                /note="SL2 trans-splice acceptor; see yk703h10.5"
gene            7533. .9207
                /gene="F35H12.5"
                /note="for a graphical representation of this gene see:
                http://www.wormbase.org/db/seq/sequence?name=F35H12.5;clas
                s=Sequence"
CDS             join(7533. .7580,7730. .7867,7922. .8227,8270. .8349,
                8659. .8744,8792. .9207)
                /gene="F35H12.5"
                /note="coded for by the following C. elegans cDNAs:
                yk703h10.5"
                /codon_start=1
                /product="Hypothetical protein F35H12.5"
                /protein_id="AAK39231.2"
                /db_xref="GI:16950403"
                /translation="MFSSRHILSNVRPIIKNLDLTRGFSDRKTKYFDLNDLIKTVNVT
                YKTGVNLEETTTVESKYVDVTPSSNTKGTVVTLSGSPGTHNDFKYMKSFFEQKKIRLI
                CTNYPGSEFVTGGLHNSYTNQDRNSYMKSLMETLELKNVNRLIIMGHSRGGENALQLT
                SMLSNDENWPLVGAVMINSPGFAPHKGISKRMGTINFIISLIKRHNKTINSILHPILH
                YFYNNLIGLRVSHGKVAAAAILPMQTFAFDEQKLSIDDLRAKPGIRAFYGYGSKDFLI
                DEHQSEEVAMYFSEEDHYVISNKQDAEKAIKEARKSFTTGKQFVTANFKEEGHFLQKT
                YPEFIVEVVDSIFDADKEVDTKI"
gene            complement(9443. .13031)
                /gene="F35H12.4"
                /note="for a graphical representation of this gene see:
                http://www.wormbase.org/db/seq/sequence?name=F35H12.4;clas
                s=Sequence"
CDS             complement(join(9443. .9627,10392. .10517,10573. .10772,
                10828. .11054,11147. .11335,11383. .11511,11615. .11845,
                11895. .12010,12526. .12728,12776. .12893,12941. .13031))
                /gene="F35H12.4"
                /note="Contains similarity to Pfam domain: PF00454
                (PI3_PI4_kinase), Score=146.8, E-value=1.6e-46, N=1; coded
                for by the following C. elegans cDNAs: yk353d4.5"
                /codon_start=1
                /product="Hypothetical protein F35H12.4"
                /protein_id="AAK39229.1"
                /db_xref="GI:13775415"
                /translation="MRLISSESFNAPMAVEYLNSIQDKPSLTFLGKRLFDLPEKDLDF
                YLPQLMYLYVTKVDAASVSHAYIEERNKKDAQFTVLCIWLLNCFTKNVPNKYENHALM
                LKSLLSRSNVPGTSQSADNMITLSLDRLEETNWTVGSSYGKSAGALARAAIQKESTLT
                AEHDSYEKSRCGPISTKLLTNQQNFIGRLVYIGVKLTEHHLLSSKQEKTIALQSELSM
                LNSMLPNSVWIPFDFENTMFNICVNESAVLNSKDKVPYVFFAEVIRSSPNYNSNIKTL
                SSVEPRIKKSQSESVLLKNKSYHDSTDPSIAVFAESWAEKKARIRGSSRYENHQKWDL
                IPVIVKTGDDLSQEAFAQQLLFTFKDLWHEEGVPLYLRPYKIVCIGSDAGLIEPVLDT
                LSLHQIKRHLTNVFRMNGNVTTPSLKHHFENVFGPVNSETYVNAQKNFIQSTAAYSLV
                SYYLQLKDRHNGNILLDMDGHLIHIDYGFLLSSSPRNLGFETAPFKLTTEIIDVMGGI
                DSDMFLYFKSLLLRGLMAARKHHRRIVSLAEIMSAGSKMQCFRAGAETVRALEARFHV
```

```
                SSTDEQLQQLVDTLVEGSRDSYTTRFYDSFQYYTNGIH"
gene            14147. .21307
                /gene="F35H12.2"
                /note="for a graphical representation of this gene see:
                http://www.wormbase.org/db/seq/sequence?name=F35H12.2a;cla
                ss=Sequence"
CDS             join(14147. .14191,14504. .14572,14623. .14699,14747. .14875,
                14935. .15001,15053. .15268,15328. .15439,15688. .15774,
                16105. .16268,18083. .18236,18290. .18492,18631. .18996,
                19415. .19558,19606. .19694,19742. .19916,20000. .20371,
                20866. .21002,21058. .21168,21220. .21307)
                /gene="F35H12.2"
                /note="coded for by the following C. elegans cDNAs:
```

```
Query Match             42.1%;  Score 631;  DB 3;  Length 27102;
Best Local Similarity   83.3%;  Pred. No. 9e-134;
Matches  822;  Conservative    0;  Mismatches   10;  Indels  155;  Gaps    3;

Qy       267 ACTATCACATCCTTTTGTCCGGGAAACAGACAGTATCGTTGAGAAGGGATTGATGTCACT 326
             |||||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5629 ACTATACACTCCTTTTGTCCGGGAAACAGACAGTATCGTTGAGAAGGGATTGATGTCACT 5688

Qy       327 TGGAAATGCTCTCGTCATGTTGTGCGTGGTCGTTCTGATGACAGTTCTGCTGATTGTTTT 386
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5689 TGGAAATGCTCTCGTCATGTTGTGCGTGGTCGTTCTGATGACAGTTCTGCTGATTGTTTT 5748

Qy       387 CTATAAATACAAGTTTTATAAGCTTATTCATGGATGGCTTATTGTCAGCAGTTTTCTTCT 446
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5749 CTATAAATACAAGTTTTATAAGCTTATTCATGGATGGCTTATTGTCAGCAGTTTTCTTCT 5808

Qy       447 TCTTTTCCTATTCACTACAATCTATGTGCA---------------------------- 476
             ||||||||||||||||||||||||||||
Db      5809 TCTTTTCCTATTCACTACAATCTATGTGCAGTAAGTTGATATATTACTATTCTCATAAAA 5868

Qy       477 ---------------AGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGT 521
                            ||||||||||||||||||||||||||||||||||||||||||||||
Db      5869 ATATCAATGTTGCAGAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGT 5928

Qy       522 TTTGTTTGGACTGGGTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCC 581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5929 TTTGTTTGGACTGGGTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCC 5988

Qy       582 ATTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTAT 641
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5989 ATTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTAT 6048

Qy       642 CAAGTACCTACCAGAATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCT 701
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6049 CAAGTACCTACCAGAATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCT 6108

Qy       702 GGTTGCCGTGCTCACACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAG 761
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6109 GGTTGCCGTGCTCACACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAG 6168

Qy       762 AAACGAGCCAATTTTCCCGGCGCTGATTTATTCGT------------------------- 796
             |||||||||||||||||||||||||||||||||
Db      6169 AAACGAGCCAATTTTCCCGGCGCTGATTTATTCGTGTAAGTTTCCTAATTTATGGAATTA 6228

Qy       797 ------------------------------------CTGGAGTCATCTATCCCTACGTTC 820
                                                 ||||||||||||||||||||||||
Db      6229 ATATTCATGACGTTTCAAATTCTAAAACATTTTCAGCTGGAGTCATCTATCCCTACGTTC 6288

Qy       821 TTGTTACTGCAGTTGAAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATA 880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6289 TTGTTACTGCAGTTGAAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATA 6348

Qy       881 ----------------------------------------------CTTCTACAGCT 891
                                                            ||||||||||||
Db      6349 GTGAGTATCACCTAAAATTTTCGAATTTTTATTCCAAAACTAATTTCAGCTTCTACAGCT 6408

Qy       892 TTTCCTGGAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATT 951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6409 TTTCCTGGAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATT 6468
```

```
Qy       952 CCTCAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTA 1011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6469 CCTCAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTA 6528

Qy      1012 AGGGTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGG 1071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6529 AGGGTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGG 6588

Qy      1072 CACGAAGAGGAAGAGAGAGGTGTGAAA 1098
             |||||||||||||||| |     |||
Db      6589 CACGAAGAGGAAGAGAGTGAGTGAAAA 6615


RESULT 4
AF038936
LOCUS       AF038936               1896 bp    mRNA    linear   MAM 16-JUN-1999
DEFINITION  Bos taurus presenilin 1 mRNA, complete cds.
ACCESSION   AF038936
VERSION     AF038936.1  GI:5070115
KEYWORDS    .
SOURCE      Bos taurus.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 1896)
  AUTHORS   Sahara,N., Shirasawa,T. and Mori,H.
  TITLE     Molecular cloning of bovine presenilin 1 gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1896)
  AUTHORS   Sahara,N., Shirasawa,T. and Mori,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1997) Molecular Biology, Tokyo Institute of
            Psychiatry, 2-1-8 Kami-Kitazawa, Setagaya-ku, Tokyo 156, Japan
FEATURES             Location/Qualifiers
     source          1. .1896
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /tissue_type="brain"
     CDS             9. .1445
                     /note="causal gene for familial Alzheimer's disease"
                     /codon_start=1
                     /product="presenilin 1"
                     /protein_id="AAD39023.1"
                     /db_xref="GI:5070116"
                     /translation="MTELPAPLSYFQNAQMSEDNHLSNTVRSQNDSRERHEHGNERRR
                     RGNTESVSNGRAPSSSQQVVEQEEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIK
                     SVSFYTRKDGQLIYTPFTEDTETVAQRALHSILNAVIMISVIVIMTILLVVLYKYRCY
                     KVIHAWLIVSSLLLLFFFSFIYLGEVFKTYNVAMDYISVALLIWNFGVVGMIAIHWKG
                     PLRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETA
                     QERNETLFPALIYSSTMVWLVNMAEGDPEAQRKVSKNSNYNAQRPANSPVTTTGTESE
                     SQDPVTESDDGGFSEEWEAQRDSRLGPHHSTAESRSAVQDLSSSILASEDPEERGVKL
                     GLGDFIFYSVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPVSI
                     TFGLIFYFATDYLVQPFMDQLAFHQFYI"
     polyA_signal    1883. .1888
BASE COUNT      499 a    439 c    452 g    506 t
ORIGIN

  Query Match           18.4%;  Score 276.2;  DB 4;  Length 1896;
  Best Local Similarity   55.1%;  Pred. No. 1.2e-52;
  Matches  659;  Conservative    0;  Mismatches  508;  Indels   28;  Gaps    5;

Qy       119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
             |||| ||| | |  | || ||   |  |||||| || ||   || || ||    |
Db       205 AAGAGGAAGAGGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAACATGTGATCATGC 264

Qy       179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
             | ||||| || ||| | || || ||||   | |||||    || || ||  |  | |  |
Db       265 TGTTTGTCCCTGTGACTCTCTGTATGGTGGTAGTTGTGGCCACCATCAAATCAGTCAGCT 324

Qy       239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
             ||||||  |  ||  |||   |||| || | |    ||| |||      || || || |

Db       325 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACTCCATTTACAGAAGACACGGAGA 381

Qy       299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
             | | |   |||  |  ||  ||| ||   |||||  |  |||||  |  | ||  | |
Db       382 CTGTGGCGCAGAGAGCCCTGCACTCAATTCTGAATGCGGTGATCATGATCAGTGTCATTG 441

Qy       359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
             |  | |||||  |||| ||| | ||| | ||||||||| ||  || ||| | || ||||
Db       442 TCATCATGACGATTCTCCTGGTGGTTCTATATAAATACAGGTGCTACAAGGTCATCCATG 501

Qy       419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
             ||||||||||         |  |  | || ||| | ||| |    || ||  ||   |
Db       502 CCTGGCTTATTGTTTCATCCCTGTTGTTGCTGTTCTTTTTCTCATTCATTTACTTGGGGG 561

Qy       479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
             ||||  | ||||  |  ||||  |      |    | | ||||  | |      | | |
Db       562 AAGTGTTTAAAACCTATAATGTTGCCATGGACTACATTTCGGTTGCGCTCCTGATCTGGA 621

Qy       539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
             | | ||| ||  | ||||||||    || ||||||||||||||| || | || || |||
Db       622 ATTTTGGTGTGGTGGGAATGATCGCCATTCATTGGAAAGGTCCACTGAGACTCCAGCAGG 681

Qy       599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
               || || ||||  ||   ||| || ||||| ||||| ||||||||||||||| |||||||
Db       682 CATATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCAGAAT 741

Qy       659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
             ||||  |  ||| |  |  ||  ||| || || || |  ||| | || || ||  |   |
Db       742 GGACCGCCTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 801

Qy       719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
             |||||||  ||| |  |     |||| ||||| || ||||||||||  ||   |  ||||
Db       802 CAAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACTCTCTTCC 861

Qy       779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
             | || || ||||| || ||  | |  | |           | || | |   |  |||
Db       862 CAGCTCTTATTTACTCTTCAACAATGGTGTGGT---------TGGTGAATATGGCAGAAG 912

Qy       839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
               ||||  |||        |  ||   | |||| |||      | ||       || |
Db       913 GAGACCCAGAAGCCCAAAGGAAGGTATCCAAAAACTCCAATTACAATGCACAAAGGCCAG 972

Qy       899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
             |  |       | |  |   |    || || |      |    | | | | | | | |
Db       973 CCAACTCACCAGTTACGACGACAGGCACAGAAAGTGAGTCACAAGACCCAGTGACAGAGA 1032

Qy       959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
             |  |   |  |  |          | |        || | | |  ||
Db      1033 GTGACGATGGTGGCTTCAGCGAAGAGTGGGAAGCCCAGAGGGACAGTCGCCTGGGGCCTC 1092

Qy      1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGA---------CGCCAATTTTCACA 1069
             | |      |||||    | || |  | || || ||         | ||   |  ||
Db      1093 ATCACTCTACAGCTGAGTCGCGATCTGCCGTTCAGGACCTTTCCAGCAGCATCCTGGCCA 1152

Qy      1070 GGCACGAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTG 1129
             | | ||    || || || || || || || |||||  | || || |||||||||||  ||
Db      1153 GTGAGGACCCAGAGGAAAGGGGAGTAAAGCTTGGATTAGGAGATTTCATTTTCTACAGTG 1212

Qy      1130 TTCTCCTCGGCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTT 1183
             ||||  | || || ||||      || |      |||||||||| || || || ||||
Db      1213 TTCTGGTTGGTAAAGCTTCTGCAACAGCCAGTGGAGACTGGAACACAACCATTGCCTGTT 1272

Qy      1184 ATGTGGCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAG 1243
             ||| |||||  | || ||| | ||| | ||  |  | || ||||||  | |||||   ||
Db      1273 TTGTAGCCATATTAATTGGTTTGTGCCTTACATTATTGCTCCTCGCCATTTTCAAGAAAG 1332

Qy      1244 CACTCCCGGCTC-TGCAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTAC 1297
             || | || |||| | || | || ||   ||  || || |||| ||||||   ||
Db      1333 CATTACCAGCTCTTCCAGTCTCTATCACCTTTGGGCTTATTTTCTACTTTGCCAC 1387


RESULT 5
DRE132931
```

```
LOCUS       DRE132931              2428 bp    mRNA    linear   VRT 05-APR-1999
DEFINITION  Danio rerio mRNA for presenilin-1.
ACCESSION   AJ132931
VERSION     AJ132931.1  GI:4583563
KEYWORDS    presenilin-1.
SOURCE      Danio rerio.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 2428)
  AUTHORS   Leimer,U. and Haass,C.
  TITLE     Expression, proteolytic processing, and amyloidogenic activity of
            zebrafish (Danio rerio) presenilin-1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2428)
  AUTHORS   Leimer,U.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-1999) Leimer U., Department of Molecular Biology,
            Central Institute of Mental Health, J5, Mannheim 68159, Germany
FEATURES             Location/Qualifiers
     source          1. .2428
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
     CDS             78. .1448
                     /codon_start=1
                     /product="presenilin-1"
                     /protein_id="CAB40386.1"
                     /db_xref="GI:4583564"
                     /translation="MADLVQNAANNVLNDGMDTSRHTSSTAAPPSRNEVELNGQPPTA
                     PPPQVVTDSEEDEDEELTLKYGAKHVIMLFIPVTLCMVVVVATIKSVSFYTQKDGQQL
                     IYTPFREDTETVGQRALHSMLNAIIMISVIVVMTLVLVVLYKYRCYKVIQAWLFFSNL
                     LLLFFFSLIYLGEVFKTYNVAMDYFTLALIIWNFGVVGMICIHWKGPLRLQQAYLIMI
                     SALMALVFIKYLPEWTAWLILAAISVYDLLAVLCPKGPLRILVETAQERNEAIFPALI
                     YSSTMVWLFNMADSAETRNNSSHPVPQQENQVVAMAPTAQAEDDGGFTPAWVDHQQHQ
                     LGPMQSTEESRRQIQEMPSARPPPPADDDEERGVKLGLGDFIFYSMLVGKASATASGD
                     WNTTLACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATDNLVRPFMDQLA
                     VHQFYI"
BASE COUNT      548 a    601 c    607 g    672 t
ORIGIN
```

Query Match 17.9%; Score 268.2; DB 5; Length 2428;
Best Local Similarity 55.1%; Pred. No. 8.1e-51;
Matches 653; Conservative 0; Mismatches 493; Indels 39; Gaps 5;

```
Qy        121 GACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATCTA 180
Db        237 GAGGAAGACGAGGACGAGGAGCTCACTCTCAAATATGGGGCGAAGCACGTCATCATGCTG 296

Qy        181 TTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGTTT 240
Db        297 TTCATTCCCGTCACGCTCTGCATGGTGGTGGTGGTGGCGACCATCAAATCCGTCAGCTTC 356

Qy        241 TATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACAGT 300
Db        357 TACACGCAGAAGGACGGACAGCAGCTGATCTACACCCCGTTTCGTGAGGACACGGAGACG 416

Qy        301 ATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCGTT 360
Db        417 GTGGGTCAGCGAGCTCTGCACTCCATGCTCAACGCCATCATCATGATCAGTGTGATTGTG 476

Qy        361 CTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATGGA 420
Db        477 GTCATGACCCTGGTGCTGGTGGTGCTCTACAAGTACAGATGCTACAAGGTGATTCAAGCC 536

Qy        421 TGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAGAA 480
Db        537 TGGCTGTTCTTCTCCAACCTCCTCCTGCTCTTCTTTTTCTCCTTAATTTATTTGGGGGAA 596

Qy        481 GTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTAAC 540
Db        597 GTGTTCAAGACGTATAACGTGGCGATGGATTACTTCACGCTGGCGTTGATCATCTGGAAC 656

Qy        541 TATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGTTC 600
Db        657 TTCGGTGTGGTGGGAATGATCTGCATCCACTGGAAGGGCCGCTGCGGCTTCAGCAGGCC 716

Qy        601 TACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAATGG 660
Db        717 TATCTGATCATGATCAGCGCTCTCATGGCTCTGGTCTTCATCAAATATCTCCCCGAGTGG 776

Qy        661 ACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACACCA 720
Db        777 ACCGCATGGCTCATCCTCGCTGCTATTTCAGTCTACGATCTTCTGGCAGTGTTGTGTCCG 836

Qy        721 AAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCCCG 780
Db        837 AAAGGCCCTCTGCGAATCCTTGTGGAAACGGCTCAAGAGAGGAATGAGGCCATTTTCCCA 896

Qy        781 GCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAAAC 840
Db        897 GCGCTCATCTACTCCTCTACGATGGTGTGGCTCTTCAATATGGCGGACAGTGCTGAAACC 956

Qy        841 ACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTGGA 900
Db        957 AGGAATAATTCCAGTCATCC----------AGTCCCTCAGCAGGA 991

Qy        901 GAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAAAA 960
Db        992 GAACCAGGTAGTTGCGATGGCTCCCACAGCTCAGGCGGAGGACGATGGCGGCTTCACTCC 1051

Qy        961 GTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGGAA 1020
Db       1052 GGCGTGGGTGGAT------CATCAGCAGCATCAGCTCGGACCCATGCAGTCCACGGAGGA 1105

Qy       1021 CGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAG-GCACGAAGA 1079
Db       1106 GAGCAGACGACAGATCCAGGAGATGCCCAGCGCTCGGCCACCACCACCAGCAGATGATGA 1165

Qy       1080 GGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCGG 1139
Db       1166 CGAAGAGCGGGGTGTGAAGCTGGGTTTGGGAGATTTCATCTTTTACAGTATGCTGGTGGG 1225

Qy       1140 CAAGGCTTCATCGTA------CTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCAT 1193
Db       1226 TAAAGCATCAGCTACAGCCAGCGGTGACTGGAACACGACGCTAGCCTGCTTCGTGGCCAT 1285

Qy       1194 TCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGGC 1253
Db       1286 TCTCATTGGTTTATGTCTGACTCTCCTCCTGCTGGCCATCTTCAAGAAGGCTCTTCCCGC 1345

Qy       1254 TCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTAC 1297
Db       1346 TCTCCCCATCTCCATAACCTTCGGCCTGGTGTTTTACTTTGCCAC 1390
```

```
RESULT 6
D84427
LOCUS       D84427                 2607 bp    mRNA    linear   VRT 06-FEB-1999
DEFINITION  Xenopus laevis mRNA for presenilin-alpha, complete cds.
ACCESSION   D84427
VERSION     D84427.1  GI:1944353
KEYWORDS    presenilin-alpha.
SOURCE      Xenopus laevis brain cDNA to mRNA.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (sites)
  AUTHORS   Tsujimura,A., Yasojima,K. and Hashimoto-Gotoh,T.
  TITLE     Cloning of Xenopus presenilin-alpha and -beta cDNAs and their
            differential expression in oogenesis and embryogenesis
  JOURNAL   Biochem. Biophys. Res. Commun. 231 (2), 392-396 (1997)
  MEDLINE   97223465
REFERENCE   2  (sites)
  AUTHORS   Tsujimura,A.
```

```
  TITLE     Isolation of Xenopus presenilin cDNAs
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2607)
  AUTHORS   Tsujimura,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-1996) Atsushi Tsujimura, Kyoto Pref.Univ.of Med.,
            Dept.Biochem. & Mol.Genet. Res.Inst.for Neuro.Dis. &Geriat.;
            Kawaramachi-Hirokoji, Kajii-Cho 465, Kamigyo-Ku, Kyoto, Kyoto
            602-0841, Japan (E-mail:atsuji@koto.kpu-m.ac.jp, Tel:075-251-5850,
            Fax:075-251-5799)
FEATURES             Location/Qualifiers
     source          1. .2607
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /tissue_type="brain"
     CDS             149. .1450
                     /codon_start=1
                     /product="presenilin-alpha"
                     /protein_id="BAA19570.1"
                     /db_xref="GI:1944354"
                     /translation="MNDTSERRSNENSESQSNGQTQSSSQQVLEQDEEEDEELTLKYG
                     AKHVIMLFVPVTLCMVVVVATIKSVSFYTRFDGQLIYTPFTEDTESVGQRALNSILNA
                     TIMISVIIVMTILLVVLYKYRCYKVIHGWLIISSLLLLFFFSYIYLGEVFKTYNVAVD
                     YITLALLIWNFGVVGMICIHWKGPLLLQQAYLIMISALMALVFIKYLPEWTTWLILAV
                     ISVYDLVAVLSPKGPLRMLVETAQERNETLFPALIYSSTMIWLVNMADGDPGLKQSAS
                     TKTYNTQAPTAHPRSDSAASDDNGGFDTTWEDHRNAQIGPINSTPESRVAVQALPSNS
                     PPSEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNTTLACFVAILIGLCLTLLLLA
                     IFKKALPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYI"
BASE COUNT      768 a    580 c    500 g    759 t
ORIGIN

  Query Match          17.6%;  Score 264.2;  DB 5;  Length 2607;
  Best Local Similarity  53.6%;  Pred. No. 6.6e-50;
  Matches  686;  Conservative    0;  Mismatches  553;  Indels   40;  Gaps    5;

Qy      116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
Db      237 AACAGGATGAGGAAGAAGATGAGGAACTCACCTTGAAATATGGAGCCAAGCATGTAATCA 296

Qy      176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTA 235
Db      297 TGCTGTTTGTCCCTGTCACACTATGCATGGTGGTGGTTGTTGCAACCATCAAGTCTGTCA 356

Qy      236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
Db      357 GCTTTTACA---CACGCTTTGATGGGCAGCTAATCTACACCCCATTCACAGAAGACACAG 413

Qy      296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
Db      414 AAAGTGTTGGACAAAGAGCCCTGAACTCCATACTCAACGCGACCATAATGATCAGTGTAA 473

Qy      356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
Db      474 TTATAGTTATGACCATACTGCTGGTTGTTCTATACAAGTACAGATGTTACAAGGTCATCC 533

Qy      416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
Db      534 ATGGCTGGCTCATCATATCTTCCCTTTTATTGCTTTTTTTCTTCTCTTATATATACCTTG 593

Qy      476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
Db      594 GTGAAGTGTTTAAAACCTATAATGTGGCCGTGGATTACATCACTCTGGCTTTACTGATCT 653

Qy      536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
Db      654 GGAATTTTGGTGTAGTAGGAATGATATGCATTCACTGGAAGGGACCCCTTCTCCTGCAGC 713

Qy      596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
Db      714 AGGCCTATCTTATCATGATTAGTGCTCTTATGGCTTTGGTGTTTATTAAATACCTTCCTG 773

Qy      656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
Db      774 AGTGGACAACTTGGCTCATCTTAGCCGTGATATCTGTTTATGATCTCGTTGCTGTCCTGT 833

Qy      716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
Db      834 CTCCAAAAGGACCTTTACGAATGCTGGTCGAGACAGCACAGGAGAGAAATGAAACCCTTT 893

Qy      776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
Db      894 TTCCAGCATTAATTTATTCTTCTACAATGATATGGTTAGTGAACATGGCAGATGGAGATC 953

Qy      836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
Db      954 CTGGTTTAAAGCAAAGTGCCTCAACGAAAACATACAACACACAAGCTCCAACTGCCCATC 1013

Qy      896 CTGGAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTC 955
Db     1014 C------GAGGAGTGATTCTGCAGCATCTGATGATAATGGTGGCTTTGATACAACTTGGG 1067

Qy      956 AAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGG 1015
Db     1068 AGGACCACAGGAATGCACAGATCGGACCAATTAATTCTACTCCAGAATCCAGAGT----- 1122

Qy     1016 TGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACG 1075
Db     1123 ------------------TGCTGTCCAGGCTCTTCCGAGCAACAGCCCCCCAAGTGAAG 1163

Qy     1076 AAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCC 1135
Db     1164 ACCCTGAAGAGAGGGGAGTGAAACTTGGACTGGGAGACTTCATTTTCTATAGTGTACTTG 1223

Qy     1136 TCGGCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGG 1189
Db     1224 TTGGAAAGGCATCTGCCACTGCTAGCGGAGACTGGAACACTACACTGGCTTGTTTTGTGG 1283

Qy     1190 CCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCC 1249
Db     1284 CCATACTGATTGGCCTTTGCCTTACATTACTTCTACTCGCCATCTTCAAGAAAGCATTAC 1343

Qy     1250 CGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCA 1308
Db     1344 CTGCTCTCCCCATCTCTATAACTTTTGGACTTGTATTTTATTTTGCCACGGACTACCTTG 1403

Qy     1309 TCACCCCATTTGTTACACAAGTCTCTCAAAAGTGTTTATTATATTAATTCTCTGTTTTTG 1368
Db     1404 TACAGCCATTCATGGACCAACTGGCATTCCACCAGTTTTATATTTAACCCACATAACTAA 1463

Qy     1369 CCATTTCTTTGCATCATCA 1387
Db     1464 CCATTCCCCGGGTTCTTAA 1482

RESULT 7
AY043492
LOCUS       AY043492                1836 bp    mRNA    linear   VRT 02-SEP-2001
DEFINITION  Gallus gallus presenilin 1 mRNA, complete cds.
ACCESSION   AY043492
VERSION     AY043492.1  GI:15421096
KEYWORDS    .
SOURCE      Gallus gallus.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1836)
  AUTHORS   Korade Mirnics,Z., Keryanov,S., Lovelock,J. and Corey,S.J.
  TITLE     Cloning of chicken presenilins
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1836)
  AUTHORS   Korade Mirnics,Z., Keryanov,S., Lovelock,J. and Corey,S.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Pediatrics, Children's Hospital of
            Pittsburgh, University of Pittsburgh, 3460.Fifth Avenue,
            Pittsburgh, PA 15213, USA
```

```
FEATURES             Location/Qualifiers
     source          1. .1836
                     /organism="Gallus gallus"
                     /isolate="E14"
                     /db_xref="taxon:9031"
                     /tissue_type="brain"
     CDS             86. .1492
                     /note="transmembrane protein; putative gamma-secretase;
                     PS1"
                     /codon_start=1
                     /product="presenilin 1"
                     /protein_id="AAK95408.1"
                     /db_xref="GI:15421097"
                     /translation="MTELSAHLPQFQHGQMTENFPDNHLSNTNDNSERRRHDNSERRR
                     NDNPGSETNGQPQNNIQQVVDQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIK
                     SVSFYTRKDGQLIYTPFTEETDTIGQRALNSILNAAIMISVIIVMTILLVVLYKYRCY
                     KVIHGWLIISSLLLLFFFSFIYLGEVFKTYNVAMDYITVALIIWNFGVVGMICIHWKG
                     PLRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETA
                     XERNETLFPALIYSSTMVWLVNMAEEDPEGQRKASKNSTYDKQAPANQSQNEDAEADD
                     GGFSQEWQQQRDNRIGPIESTPESRAAVQALPSNSQTSEDPEERGVKLGLGDFIFYSV
                     LVGKASATASGDWNTTLACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFAT
                     DNLVQPFMDQLAFHQFYI"
BASE COUNT      502 a    389 c    406 g    529 t     10 others
ORIGIN
```

Query Match 17.4%; Score 261.6; DB 5; Length 1836;
Best Local Similarity 55.2%; Pred. No. 2.6e-49;
Matches 655; Conservative 0; Mismatches 491; Indels 40; Gaps 6;

```
Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||| | |  |  || ||       ||||||| || |||   || || |||
Db        282 AAGATGAAGAAGAAGATGAGGAGCTAACACTGAAATATGGGGCAAAACATGTGATTATGT 341

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||  |||| |||||||  | ||||||  || || ||  |  | |  |
Db        342 TGTTTGTACCTGTCACACTTTGCATGGTGGTCGTTGTTGCAACCATCAAGTCTGTCAGCT 401

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
              ||||||  |  ||  |||||  ||  |  | |   ||| ||       || ||||| |
Db        402 TTTATACACGCAAGGATGGA---CAGCTTATCTACACTCCCTTCACAGAAGAGACAGATA 458

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
              || |   ||| ||  |||  ||| ||  || |||    ||||| |  | ||  ||
Db        459 CCATTGGACAGAGGGCCCTGAATTCAATTCTCAACGCTGCAATCATGATTAGTGTCATCA 518

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
              || | |||||  | ||  ||| |||  | ||  |||||||  ||  || ||| | ||  || |
Db        519 TTGTCATGACCATACTCCTGGTTGTGCTTTACAAATACAGGTGCTACAAGGTGATCCACG 578

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
              | ||||||||  |      | ||  | ||  |||||| | ||  |    ||||||  |    |
Db        579 GTTGGCTTATCATTTCTTCTCTTTTGCTGCTTTTCTTTTTTTCATTCATCTATCTAGGTG 638

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
              | ||| | || |  |  | ||  |       |    ||      |   | | |  | |
Db        639 AGGTTTTTAAGACGTATAACGTTGCCATGGACTACATTACAGTGGCACTCATAATCTGGA 698

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
              ||| ||| ||| | |||||||| |||||  |||||||| || ||| | ||  || || ||
Db        699 ACTTTGGTGTTGTGGGAATGATTTGTATCCATTGGAAGGGCCCACTTCGGCTTCAGCAAG 758

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
              || || ||||  ||    || || |||||  |||| || || |||||  || ||  ||||
Db        759 CATATCTCATTATGATAAGCGCTCTCATGGCATTGGTGTTCATTAAGTATCTTCCCGAAT 818

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
              ||||||  ||| || | |||  ||| || || || |  |||  ||||||| ||  |    |
Db        819 GGACTGCATGGCTTATCCTGGCTGTCATTTCAGTGTACGATTTGGTTGCTGTCTTATGTC 878

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
              | ||||| ||  |  |    || || ||||| || || |||||||| ||  |  | ||||
Db        879 CTAAAGGTCCTCTTCGTATGTTAGTTGAAACAGCTCANGAGAGAAATGAAACGCTCTTCC 938

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
              | || || ||||| || ||   | |  | |           || || |     | |||
Db        939 CAGCACTTATTTACTCCTCAACAATGGTGTGG---------CTAGTGAACATGGCTGAGG 989

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
              |     ||||  ||         | ||   | |||| | |     || |     ||| |
Db        990 AAGATCCAGAAGGCCAACGGAAAGCTTCTAAAAACTCCACTTACGATAAACAAGCTCCGG 1049

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
              | |   |  |          ||  |   |       || | |     |  ||  ||||
Db        1050 CAAACCAGTCTCAGAATGAAGATGCTGAA------GCAGATGATGGNGGCTTTAGTCAAG 1103

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
              |   |||    .||    | | || |     | | |   ||    |
Db        1104 AGTGGCAGCAACAAAGAGACAATAGAATAGGACCCATTGAATCAAC------------- 1149

Qy        1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
               ||  ||   |  ||||     |   || | |       | |    |||  | ||
Db        1150 -ACCTGAATCACGAGCTGCTGTTCAGGCTCTGCCTAGTAACTCTCAAACCAGTGAAGACC 1208

Qy        1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
               ||||| ||||| || || || ||| | || ||||||||||||||||   || ||  | |
Db        1209 CTGAAGAAAGAGGAGTAAAGCTAGGTTTAGGAGACTTCATTTTCTACAGCGTCCTGGTTG 1268

Qy        1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
               |||| || |      |  |   || ||||| || ||| | || |||| ||| ||||
Db        1269 GCAAAGCCTCTGCAACGGCAAGCGGGGATTGGAATACAACTTTAGCATGTTTTGTAGCCA 1328

Qy        1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
               |  | || ||||| ||| | || ||| | || || ||| |||| ||    ||  | || |
Db        1329 TATTAATTGGTCTGTGCCTTACACTTCTGCTTCTTGCCATCTTTAAGAAGGCCTTACCAG 1388

Qy        1253 CTC-TGCAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTAC 1297
               ||| | |||| ||||||   |  || ||  ||||||||||  |||
Db        1389 CTCTTCCAATCTCCATTACTTTTGGGCTAGTTTTTTACTTTGCTAC 1434
```

```
RESULT 8
MMPRSNLNI
LOCUS       MMPRSNLNI               1430 bp    mRNA    linear   PRI 19-DEC-1996
DEFINITION  M.murinus presenilin I gene.
ACCESSION   Z71333
VERSION     Z71333.1  GI:1707591
KEYWORDS    presenilin I.
SOURCE      Microcebus murinus.
  ORGANISM  Microcebus murinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
            Microcebus.
REFERENCE   1  (bases 1 to 1430)
  AUTHORS   Calenda,A., Mestre-Frances,N., Czech,C., Pradier,L., Petter,A.,
            Bons,N. and Bellis,M.
  TITLE     Molecular cloning, sequencing, and brain expression of the
            presenilin 1 gene in Microcebus murinus
  JOURNAL   Biochem. Biophys. Res. Commun. 228 (2), 430-439 (1996)
  MEDLINE   97079199
   PUBMED   8920931
REFERENCE   2  (bases 1 to 1430)
  AUTHORS   Calenda,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-1996) Calenda A., INSERMU249/CNRS UPR9008/CRBM,
            Institut de biologie, boulevard Henri IV, Montpellier, FRANCE,
            34060
FEATURES             Location/Qualifiers
     source          1. .1430
                     /organism="Microcebus murinus"
                     /db_xref="taxon:30608"
                     /clone="ST2884"
                     /sex="Male"
                     /tissue_type="lemurian brain"
     CDS             1. .1404
                     /codon_start=1
                     /product="presenilin I"
```

```
                 /protein_id="CAA95930.1"
                 /db_xref="GI:1707592"
                 /db_xref="SPTREMBL:P79802"
                 /translation="MTELPAPLSYFQNAQMSEDNHLSNTVRSQNDNREQQDHGDRRRL
                 GNPEPLSNGRPQGNSGPVVERDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                 VSFYTRKDGQLIYTPFTEDTETVGQRALHSVLNAAIMISVIVVMTILLVVLYKYRCYK
                 VIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGP
                 LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                 ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNTKYNAQGTEREAQASVPENDDG
                 GFSEEWEAQRDSQLGPHRSTSVSRAAVQETSSSIPASEDPEERGVKLGLGDFVFYSVL
                 VGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATD
                 YLVQPFMDQLAFHQFYI"
BASE COUNT      373 a    324 c    339 g    394 t
ORIGIN

  Query Match           17.4%;  Score 261;  DB 9;  Length 1430;
  Best Local Similarity  55.3%;  Pred. No. 3.6e-49;
  Matches  655;  Conservative    0;  Mismatches  490;  Indels   40;  Gaps    6;

Qy       120 AGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATCT 179
Db       195 AGATGAGGAAGAAGATGAGGAGCTGACATTGAAATACGGCGCCAAGCATGTGATCATGCT 254

Qy       180 ATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGTT 239
Db       255 CTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTTGTGGCTACCATCAAATCGGTCAGCTT 314

Qy       240 TTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACAG 299
Db       315 TTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGACACCGAGAC 371

Qy       300 TATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCGT 359
Db       372 TGTGGGCCAGAGAGCCCTGCACTCAGTTCTGAATGCTGCCATCATGATCAGTGTCATTGT 431

Qy       360 TCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATGG 419
Db       432 TGTCATGACTATCCTCCTGGTGGTTCTATATAAATATAGGTGCTATAAGGTCATCCATGC 491

Qy       420 ATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAGA 479
Db       492 CTGGCTGATCATATCATCTCTATTGTTGCTGTTCTTTTTTTCCTTCATTTACTTGGGAGA 551

Qy       480 AGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTAA 539
Db       552 AGTGTTTAAAACCTATAATGTTGCTGTGGACTACATTACTGTTGCGCTCCTGATCTGGAA 611

Qy       540 CTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGTT 599
Db       612 TTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAACAGGC 671

Qy       600 CTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAATG 659
Db       672 ATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAATG 731

Qy       660 GACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACACC 719
Db       732 GACTGCGTGGCTCATCTTGGCTGTAATCTCAGTATATGATTTAGTGGCTGTTTTGTGTCC 791

Qy       720 AAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCCC 779
Db       792 AAAAGGTCCACTTCGTATGCTGGTGGAAACAGCTCAGGAAAGAAATGAAACTCTATTTCC 851

Qy       780 GGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAAA 839
Db       852 AGCTCTCATTTACTCCTCAACAATGGTGTGGT---------TGGTGAATATGGCAGAAGG 902

Qy       840 CACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTGG 899
Db       903 AGACCCAGAAGCCCAAAGGAGGGTATCCAAAAACACTAAGTATAATGCACAAGGTACAGA 962

Qy       900 AGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAAA 959

Db       963 AAGGGAGGCACAAGCTTCTG-TTCCAGAGAATGATGATGGCGGCTTCAGTGAAGAGTGGG 1021

Qy       960 AGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGGA 1019
Db      1022 AAGCCCAAAGGGACAGTCAACTAGGGCCTCATCGCTCTACATCT--------------- 1065

Qy      1020 ACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAGA 1079
Db      1066 ----GTGTCACGAGCTGCTGTCCAGGAAATTTCCAGCAGCATCCCAGCCAGTGAAGATCC 1121

Qy      1080 GGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCGG 1139
Db      1122 AGAGGAAAGGGGAGTAAAACTTGGATTAGGAGATTTCGTTTTCTACAGTGTTCTGGTTGG 1181

Qy      1140 CAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCAT 1193
Db      1182 CAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTTGTAGCCAT 1241

Qy      1194 TCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGGC 1253
Db      1242 ATTAATTGGCTTGTGCCTTACATTACTACTCCTCGCCATTTTCAAGAAAGCATTGCCAGC 1301

Qy      1254 TC-TGCAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTAC 1297
Db      1302 TCTTCCAATCTCTATCACCTTCGGGCTTGTTTTCTACTTTGCCAC 1346

RESULT 9
RATS182
LOCUS       RATS182                 1407 bp    mRNA    linear   ROD 11-JAN-2002
DEFINITION  Rattus norvegicus mRNA for S182 protein, complete cds.
ACCESSION   D82578
VERSION     D82578.1  GI:1139562
KEYWORDS    .
SOURCE      Rattus norvegicus (strain:Wistar) brain cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Taniguchi,T., Hashimoto,T., Taniguchi,R., Shimada,K., Kawamata,T.,
            Yasuda,M., Nakai,M., Terashima,A., Koizumi,T., Maeda,K. and
            Tanaka,C.
  TITLE     Cloning of the cDNA encoding rat Presenilin-1
  JOURNAL   Gene 186 (1), 73-75 (1997)
  MEDLINE   97199371
REFERENCE   2  (bases 1 to 1407)
  AUTHORS   Taniguchi,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-1995) Taizo Taniguchi, Hyogo Institute for Aging
            Brain and Cognitive Disorders, Department of Neuroscience; 520
            Saisho-ko, Himeji, Hyogo 670, Japan (E-mail:tanigu@hiabcd.go.jp,
            Tel:0792-95-5511, Fax:0792-95-8199)
FEATURES             Location/Qualifiers
     source          1. .1407
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /tissue_type="brain"
     CDS             1. .1407
                     /note="presenilin 1"
                     /codon_start=1
                     /product="S182 protein"
                     /protein_id="BAA11575.1"
                     /db_xref="GI:1783304"
                     /translation="MTEIPAPLSYFQNAQMSEDSHSSSVRSQNDNQERQQHHDRQRLD
                     NPESISNGRPQSNFTRQVIEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                     VSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYK
                     VIHAWLIVSSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMIAIHWKGP
                     LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                     ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVPKNPKYSTQGTEREETQDTGTGSDD
                     GGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSILTSEDPEERGVKLGLGDFIFYSV
                     LVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLIFYFAT
```

```
                DYLVQPFMDQLAFHQFYI"
BASE COUNT      346 a    357 c    348 g    356 t
ORIGIN

  Query Match           17.2%;  Score 258.4;  DB 10;  Length 1407;
  Best Local Similarity   53.8%;  Pred. No. 1.4e-48;
  Matches  667;  Conservative    0;  Mismatches  536;  Indels   37;  Gaps    5;

Qy         98 TAACAAATCGGAATAGCCAAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACG 157
              | || |  | |    |   |  |||| ||  | |  |  |||||   |    ||||||| |
Db        173 TTACCAGGCAGGTGATAGAGCAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATG 232

Qy        158 GAGCATCTCACGTTATTCATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTT 217
              ||||    |||||  ||    || ||||| || ||| | || |||||||   | || ||
Db        233 GAGCCAAGCACGTCATCATGCTCTTTGTTCCTGTGACCCTCTGCATGGTCGTTGTGGTGG 292

Qy        218 TTACGATGAACACGATTACGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATC 277
                || || ||  |  | |  || || |  |  ||  ||| . ||||  || | |     |
Db        293 CCACCATCAAGTCAGTCAGCTTCTACACCCGGAAGGATG---GGCAGCTAATCTATACCC 349

Qy        278 CTTTTGTCCGGGAAACAGACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTC 337
              | ||       || || || | | | |   |||  |   ||   ||  |    |||||
Db        350 CATTCACAGAAGACACCGAGACTGTAGGCCAGAGAGCCCTGCACTCGATCTTGAATGCCG 409

Qy        338 TCGTCATGTTGTGCGTGGTCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACA 397
              | ||||| |  | ||  |||| | ||||  | || ||| | ||  | ||||| ||||
Db        410 CCATCATGATCAGTGTCATCGTCGTTATGACCATCCTCCTGGTGGTCCTGTATAAGTACA 469

Qy        398 AGTTTTATAAGCTTATTCATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTAT 457
              ||  || ||| | || ||||  ||||||||||       | |  |  | || ||| | |
Db        470 GGTGCTACAAGGTCATCCATGCCTGGCTTATTGTTTCATCTCTGTTGTTGCTGTTCTTTT 529

Qy        458 TCACTACAATCTATGTGCAAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTAT 517
              |  |    || ||  |    |||||  | || |  | | ||||  |       |    |
Db        530 TTTCATTCATTTACTTAGGGGAAGTATTCAAGACCTACAATGTCGCCGTGGACTATATTA 589

Qy        518 TGGTTTTGTTTGGACTGGGTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAG 577
              ||||    |     |  | || | ||| ||  |||| |||||     || || |||||||
Db        590 CGGTTGCACTCCTGATCTGGAATTTTGGTGTGGTCGGGATGATTGCCATTCACTGGAAAG 649

Qy        578 GTCCATTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCT 637
              | ||| | || ||||| |||   || || ||||  ||  ||| || ||||| ||||| |
Db        650 GCCCACTCCGACTGCAGCAGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTAT 709

Qy        638 TTATCAAGTACCTACCAGAATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGG 697
              |||||||||||||| || |||||||| |  ||| |  |  ||  ||| || || || |  |
Db        710 TTATCAAGTACCTCCCTGAATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATG 769

Qy        698 ATCTGGTTGCCGTGCTCACACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGG 757
              || |||| || || ||    || ||||| ||| .|  |     |||| ||||| || ||||
Db        770 ATTTGGTGGCTGTTCTGTGCCCCAAAGGTCCACTTCGTATGCTGGTCGAAACAGCTCAGG 829

Qy        758 AGAGAAACGAGCCAATTTTCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACG 817
              | ||||| ||| |  | || || || || || ||||| ||    |  | |
Db        830 AAAGAAATGAGACTCTCTTTCCAGCTCTTATCTATTCCTCAACCATGGTGTGGTTGGTGA 889

Qy        818 TTCTTGTTACTGCAGTTGAAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAA 877
              | | | |                ||  ||  |||   | || |         || |||
Db        890 ATATGGCT----------------GAAGGAGACCCAGAAGCCCAAAGGAGGGTACCCAA 932

Qy        878 ATACTTCTACAGCTTTTCCTGGAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAA 937
              | ||  | |         |     ||   |||  |         | | | ||      |
Db        933 AAACCCCAAGTATAGCACACAAGGAACAGAGAGGGAAGAGACACAGGACACTGGCACTGG 992

Qy        938 AAGTGAAACGAATTCCTCAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACAC 997
              || ||       |  ||  |     |           | ||    ||||  |   |   |
Db        993 GAGCGATGATGGTGGCTTCAGTGAGGAGTGGGAGGCCCAAAGAGACAGTCACCTGGGGCC 1052

Qy        998 AAAACTCTGGAGTAAGGGTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACG 1057
              | | ||  | |           |  |||  |  ||||      |       |     |
Db       1053 TCATCGCTCCACT----------CCTGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGA 1102

Qy       1058 CCAATTTTCACAGGCACGAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCA 1117
              ||   |    ||  | ||   ||| || ||||| || || ||||| ||||| || ||||
Db       1103 GCATCCTCACTAGTGAAGACCCGGAGGAAAGAGGAGTAAAGCTTGGGCTGGGAGATTTCA 1162

Qy       1118 TTTTCTACTCTGTTCTCCTCGGCAAGGCTTCATCGTAC------TTTGACTGGAACACGA 1171
              |||||||  ||||||  | || ||||| ||| ||  |        ||||||||||| |
Db       1163 TTTTCTACAGTGTTCTGGTTGGTAAGGCCTCAGCGACCGCCAGTGGGGACTGGAACACAA 1222

Qy       1172 CTATCGCTTGTTATGTGGCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCG 1231
              | || || || | ||| |||||  | ||||| || ||| | ||  |  ||||||||||
Db       1223 CCATAGCCTGCTTTGTAGCCATATTGATCGGCCTGTGCCTTACGTTACTCCTGCTCGCCA 1282

Qy       1232 TCTTCAAACGAGCACTCCCGGCTC-TGCAATTTCCATTTTCTCCGGACTCATTTTTTACT 1290
              | |||||   |||  | ||||| | | | || |||||   || ||| |||||||| ||||
Db       1283 TTTTCAAGAAAGCGTTGCCGGCCCTTCCCATCTCCATCACCTTCGGGCTCATTTTCTACT 1342

Qy       1291 TTTGTACCCGCTGGATCATCACCCCATTTGTTACACAAGT 1330
              ||   ||    |   || |    || ||  |    ||| |
Db       1343 TTGCCACGGATTATCTCGTGCAGCCCTTCATGGACCAACT 1382


RESULT 10
D82363
LOCUS       D82363                  1407 bp    mRNA    linear   ROD 06-FEB-1999
DEFINITION  Rattus norvegicus mRNA for presenilin-1, complete cds.
ACCESSION   D82363
VERSION     D82363.1  GI:1777325
KEYWORDS    presenilin 1.
SOURCE      Rattus norvegicus (strain:Wistar) embryo male brain cDNA to mRNA,
            clone:rPS-1.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Takahashi,H., Murayama,M., Takashima,A., Mercken,M., Nakazato,Y.,
            Noguchi,K. and Imahori,K.
  TITLE     Molecular cloning and expression of the rat homologue of
            presenilin-1
  JOURNAL   Neurosci. Lett. 206 (2-3), 113-116 (1996)
  MEDLINE   96255262
REFERENCE   2  (bases 1 to 1407)
  AUTHORS   Takahashi,H., Takashima,A.A. and Murayama,M.M.
  TITLE     Molecular cloning and expression of rat presenilin-1
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1407)
  AUTHORS   Takahashi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-1995) Hiroshi Takahashi, Mitsubishi Kasei
            Institute of Life Sciences, Laboratory of Histochemistry; 11
            Minamiooya, Machida, Tokyo 194, Japan (Tel:0427-24-6211,
            Fax:0427-24-6314)
FEATURES             Location/Qualifiers
     source          1. .1407
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /clone="rPS-1"
                     /sex="male"
                     /tissue_type="brain"
                     /dev_stage="embryo"
     CDS             1. .1407
                     /codon_start=1
                     /product="presenilin-1"
                     /protein_id="BAA11564.1"
                     /db_xref="GI:1777326"
                     /translation="MTEIPAPLSYFQNAQMSEDSHSSSVRSQNDNQERQQHHDRQRLD
                     NPESISNGRPQSNFTRQVIEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                     VSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYK
                     VIHAWLIVSSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMIAIHWKGP
                     LRLQQAYLIMISALMSLVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                     ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVPKNPKYSTQGTEREETQDTGTGSDD
                     GGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSILTSEDPEERGVRLGLGDFIFYSV
                     LVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLIFYFAT
```

```
                    DYLVQPFMDQLAFHQFYI"
BASE COUNT      345 a    356 c    348 g    358 t
ORIGIN

  Query Match           17.0%;  Score 255.2;  DB 10;  Length 1407;
  Best Local Similarity  53.6%;  Pred. No. 7.6e-48;
  Matches  665;  Conservative    0;  Mismatches  538;  Indels   37;  Gaps    5;

Qy         98 TAACAAATCGGAATAGCCAAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACG 157
Db        173 TTACCAGGCAGGTGATAGAGCAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATG 232

Qy        158 GAGCATCTCACGTTATTCATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTT 217
Db        233 GAGCCAAGCACGTCATCATGCTCTTTGTTCCTGTGACCCTCTGCATGGTCGTTGTGGTGG 292

Qy        218 TTACGATGAACACGATTACGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATC 277
Db        293 CCACTATCAAGTCAGTCAGCTTCTACACCCGGAAGGATG---GGCAGCTAATCTATACCC 349

Qy        278 CTTTTGTCCGGGAAACAGACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTC 337
Db        350 CATTCACAGAAGACACCGAGACTGTAGGCCAGAGAGCCCTGCACTCGATCTTGAATGCCG 409

Qy        338 TCGTCATGTTGTGCGTGGTCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACA 397
Db        410 CCATCATGATCAGTGTCATCGTCGTTATGACCATCCTCCTGGTGGTCCTGTATAAGTACA 469

Qy        398 AGTTTTATAAGCTTATTCATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTAT 457
Db        470 GGTGCTACAAGGTCATCCATGCCTGGCTTATTGTTTCATCTCTGTTGTTGCTGTTCTTTT 529

Qy        458 TCACTACAATCTATGTGCAAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTAT 517
Db        530 TTTCATTCATTTACTTAGGGGAAGTATTCAAGACCTACAATGTCGCCGTGGACTATATTA 589

Qy        518 TGGTTTTGTTTGGACTGGGTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAG 577
Db        590 CGGTTGCACTCCTGATCTGGAATTTTGGTGTGGTCGGGATGATTGCCATTCACTGGAAAG 649

Qy        578 GTCCATTGCGTCTGCAACAGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCT 637
Db        650 GCCCACTCCGACTGCAGCAGGCGTATCTCATTATGATCAGTGCCCTCATGTCCCTGGTAT 709

Qy        638 TTATCAAGTACCTACCAGAATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGG 697
Db        710 TTATCAAGTACCTCCCTGAATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATG 769

Qy        698 ATCTGGTTGCCGTGCTCACACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGG 757
Db        770 ATTTGGTGGCTGTTCTGTGCCCCAAAGGTCCACTTCGTATGCTGGTCGAAACAGCTCAGG 829

Qy        758 AGAGAAACGAGCCAATTTTCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACG 817
Db        830 AAAGAAATGAGACTCTCTTTCCAGCTCTTATCTATTCCTCAACCATGGTGTGGTTGGTGA 889

Qy        818 TTCTTGTTACTGCAGTTGAAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAA 877
Db        890 ATATGGCT----------------GAAGGAGACCCAGAAGCCCAAAGGAGGGTACCCAA 932

Qy        878 ATACTTCTACAGCTTTTCCTGGAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAA 937
Db        933 AAACCCCAAGTATAGCACACAAGGAACAGAGAGGGAAGAGACACAGGACACTGGCACTGG 992

Qy        938 AAGTGAAACGAATTCCTCAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACAC 997
Db        993 GAGCGATGATGGTGGCTTCAGTGAGGAGTGGGAGGCCCAAAGAGACAGTCACCTGGGGCC 1052

Qy        998 AAAACTCTGGAGTAAGGGTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACG 1057
Db       1053 TCATCGCTCCACT----------CCTGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGA 1102

Qy       1058 CCAATTTTCACAGGCACGAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCA 1117

Db       1103 GCATCCTCACTAGTGAAGACCCGGAGGAAAGAGGAGTAAGGCTTGGGCTGGGAGATTTCA 1162

Qy       1118 TTTTCTACTCTGTTCTCCTCGGCAAGGCTTCATCGTAC------TTGACTGGAACACGA 1171
Db       1163 TTTTCTACAGTGTTCTGGTTGGTAAGGCCTCAGCGACCGCCAGTGGGGACTGGAACACAA 1222

Qy       1172 CTATCGCTTGTTATGTGGCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCG 1231
Db       1223 CCATAGCCTGCTTTGTAGCCATATTGATCGGCCTGTGCCTTACGTTACTCCTGCTCGCCA 1282

Qy       1232 TCTTCAAACGAGCACTCCCGGCTC-TGCAATTTCCATTTTCTCCGGACTCATTTTTTACT 1290
Db       1283 TTTTCAAGAAAGCGTTGCCGGCCCTTCCCATCTCCATCACCTTCGGGCTCATTTTCTACT 1342

Qy       1291 TTTGTACCCGCTGGATCATCACCCCATTTGTTACACAAGT 1330
Db       1343 TTGCCACGGATTATCTCGTGCAGCCCTTCATGGACCAACT 1382


RESULT 11
CCPS1
LOCUS       CCPS1                     1241 bp    mRNA    linear   VRT 25-AUG-1998
DEFINITION  Cyprinus carpio mRNA for PS1 protein.
ACCESSION   Y17128
VERSION     Y17128.1  GI:3087753
KEYWORDS    PS1 gene.
SOURCE      common carp.
  ORGANISM  Cyprinus carpio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Cyprinus.
REFERENCE   1  (bases 1 to 1241)
  AUTHORS   Archer,S., Hirano,J., Diss,J.K., Fraser,S.P. and Djamgoz,M.B.
  TITLE     Expression and localization in the fish retina of a homologue of
            the Alzheimer's related PS1 gene
  JOURNAL   Neuroreport 9 (9), 2049-2056 (1998)
  MEDLINE   98337349
   PUBMED   9674592
REFERENCE   2  (bases 1 to 1241)
  AUTHORS   Archer,S.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-APR-1998) S.N. Archer, International Marine Centre,
            Localita sa Mardini, 09072 Torregrande, Oristano, ITALY
FEATURES             Location/Qualifiers
     source          1. .1241
                     /organism="Cyprinus carpio"
                     /db_xref="taxon:7962"
                     /dev_stage="adult"
     gene            1. .1156
                     /gene="PS1"
     CDS             <1. .1156
                     /gene="PS1"
                     /codon_start=2
                     /product="PS1 protein"
                     /protein_id="CAA76641.1"
                     /db_xref="GI:3087754"
                     /db_xref="SPTREMBL:O73869"
                     /translation="IMLFIPVTLCMVVVVATIKSVSFYTQKDGQQLIYTPFREDTETV
                     GQRALNSMLNAAIMISVIVVMTLVLVVLYKYRCYKVIQGWLFFSNLLLLFFFSFIYLG
                     EVFKTYNVAMDYFTLAVIIWNFGVVGMICIHWKGPLRLQQAYLIMISALMALVFIKYL
                     PEWTAWLILAAISVYDLLAVLCPKGPLRILVETAQERNEPIFPALIYSSTMVWLFNMA
                     DSAEPRNNSSHPVPQQENQDAVAPTAQPEDDGGFTPAWVNQQQHQLGPMQSTEDSRRE
                     IQELPSARPPPVEDDEERGVKLGLGDFIFYSMLVGKASATASGDWNTTIACFVAILIG
                     LCLTLLLLAIFKKALPALPISITFGLVFYFATDNLVRPFMDQLAVHQFYI"
     variation       14
                     /gene="PS1"
                     /note="variant"
                     /replace="g"
     variation       76
                     /gene="PS1"
                     /note="variant"
                     /replace="t"
```

```
variation       136
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       142
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       145
                /gene="PS1"
                /note="variant"
                /replace="a"
variation       148
                /gene="PS1"
                /note="variant"
                /replace="g"
variation       167
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       168
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       179
                /gene="PS1"
                /note="variant"
                /replace="g"
variation       181
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       188
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       249
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       271
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       286
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       295
                /gene="PS1"
                /note="variant"
                /replace="a"
variation       322
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       353
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       454
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       532
                /gene="PS1"
                /note="variant"
                /replace="a"
variation       538
                /gene="PS1"
                /note="variant"
                /replace="a"
variation       559
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       560
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       571
                /gene="PS1"
                /note="variant"
                /replace="g"
variation       580
                /gene="PS1"
                /note="variant"
                /replace="a"
variation       599
                /gene="PS1"
                /note="variant"
                /replace="a"
variation       839
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       850
                /gene="PS1"
                /note="variant"
                /replace="t"
variation       862
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       874
                /gene="PS1"
                /note="variant"
                /replace="a"
variation       946
                /gene="PS1"
                /note="variant"
                /replace="c"
variation       949
                /gene="PS1"
                /note="variant"
                /replace="g"
BASE COUNT      257 a    348 c    323 g    313 t
ORIGIN
```

Query Match 16.8%; Score 252.2; DB 5; Length 1241;
Best Local Similarity 53.5%; Pred. No. 3.7e-47;
Matches 668; Conservative 0; Mismatches 538; Indels 43; Gaps 5;

```
Qy      178 CTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACG 237
Db        8 CTGTTTATTCCCGTCACGCTGTGTATGGTGGTGGTTGTGGCCACCATCAAATCGGTCAGT 67

Qy      238 TTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGAC 297
Db       68 TTCTACACGCAGAAGGACGGCCAGCAGCTGATCTACACTCCGTTCCGTGAGGACACAGAG 127

Qy      298 AGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTC 357
Db      128 ACGGTGGGGCAGCGAGCGCTCAACTCCATGCTCAACGCCGCCATCATGATCAGTGTGATT 187

Qy      358 GTTCTGATGACACGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCAT 417
Db      188 GTGGTCATGACCCTGGTGCTGGTGGTGCTCTACAAGTACAGATGCTACAAGGTGATTCAA 247

Qy      418 GGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAA 477
Db      248 GGCTGGCTGTTCTTCTCCAACCTTCTCCTCCTCTTCTTTTTCTCCTTCATTTATTTGGGG 307

Qy      478 GAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGT 537
Db      308 GAAGTGTTCAAGACGTACAACGTGGCCATGGACTACTTCACGTTAGCCGTGATCATCTGG 367
```

```
Qy        538 AACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAG 597
Db        368 AACTTCGGTGTGGTGGGAATGATCTGTATCCACTGGAAGGGTCCGCTGCGGCTCCAGCAG 427

Qy        598 TTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAA 657
Db        428 GCCTATCTGATCATGATCAGCGCTCTAATGGCCCTGGTCTTCATCAAGTACCTGCCGGAG 487

Qy        658 TGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACA 717
Db        488 TGGACCGCCTGGCTCATCCTCGCTGCTATATCTGTCTACGATCTGTTGGCCGTGTTGTGC 547

Qy        718 CCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTC 777
Db        548 CCGAAAGGACCCTTGCGTATCCTTGTGGAAACGGCTCAAGAGAGGAATGAACCCATCTTC 607

Qy        778 CCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAA 837
Db        608 CCTGCTCTCATCTACTCCTCAACCATGGTGTGGCTGTTTAATATGGCGGACAGCGCAGAG 667

Qy        838 AACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCT 897
Db        668 CCCAGGAATAATTCCAGCCA----------------------------TCCAGTTCCTCAG 700

Qy        898 GGAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAA 957
Db        701 CAGGAGAACCAGGATGCCGTGGCACCCACGGCTCAGCCGGAGGATGATGGCGGTTTCACA 760

Qy        958 AAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTG 1017
Db        761 CCAGCGTGGGTC-------AATCAGCAGCAGCATCAGCTCGGCCCCATGCAGTCCACTGA 813

Qy        1018 GAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAA 1077
Db         814 GGACAGCAGACGAGAGATCCAAGAGCTGCCCTCC--GCACGCCCGCCACCTGTGGAAGAC 871

Qy        1078 GAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTC 1137
Db         872 GACGAAGAGCGGGGTGTGAAGCTGGGTTTGGGAGATTTCATCTTTTACAGTATGCTGGTG 931

Qy        1138 GGCAA------GGCTTCATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCC 1191
Db         932 GGTAAAGCATCAGCTACAGCCAGCGGTGACTGGAACACGACGATAGCCTGCTTTGTGGCC 991

Qy        1192 ATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCG 1251
Db         992 ATTCTCATCGGTTTATGTCTAACTCTGCTCCTCCTGGCGATCTTCAAAAAGGCCCTTCCG 1051

Qy        1252 GCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATC 1310
Db        1052 GCTCTCCCCATCTCCATAACCTTCGGCCTGGTCTTTTACTTTGCCACCGATAACCTCGTG 1111

Qy        1311 ACCCCATTTGTTACACAAGTCTCTCAAAAGTGTTTATTATATTAATTCTCTGTTTTTGCC 1370
Db        1112 CGGCCGTTCATGGATCAGCTGGCTGTTCATCAGTTCTACATTTAGCACGGAGGAGAGGTC 1171

Qy        1371 ATTTCTTTGCATCATCAACTTTTCGATTATATCTTGAGCGATCTCAAAG 1419
Db        1172 ACTCCTCGTGCGCTTCACTTTACAGCTAATTCCCAGAATCCTCTTGCAG 1220
```

RESULT 12
HSU40379

```
LOCUS       HSU40379                1392 bp    mRNA    linear   PRI 12-AUG-1996
DEFINITION  Human presenilin I-463 (AD3-3) mRNA, complete cds.
ACCESSION   U40379
VERSION     U40379.1  GI:1244637
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1392)
  AUTHORS   Sahara,N., Yahagi,Y., Takagi,H., Kondo,T., Okochi,M., Usami,M.,
            Shirasawa,T. and Mori,H.
  TITLE     Identification and characterization of presenilin I-467, I-463 and
            I-374
  JOURNAL   FEBS Lett. 381 (1-2), 7-11 (1996)
  MEDLINE   96193901
   PUBMED   8641442
REFERENCE   2  (bases 1 to 1392)
  AUTHORS   Shirasawa,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-1995) Takuji Shirasawa, Molecular Pathology,
            Tokyo Metropolitan Institute of Gerontology, 35-2 Sakae-cho,
            Itabashi-ku, Tokyo, 175, Japan
FEATURES             Location/Qualifiers
     source          1. .1392
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q24.3"
                     /clone="AD3-3"
     gene            1. .1392
                     /gene="AD3-3"
     CDS             1. .1392
                     /gene="AD3-3"
                     /codon_start=1
                     /product="presenilin I-463"
                     /protein_id="AAB05894.1"
                     /db_xref="GI:1244638"
                     /translation="MTELPAPLSYFQNAQMSEDNHLSNTNDNRERQEHNDRRSLGHPE
                     PLSNGRPQGNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFY
                     TRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHA
                     WLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQ
                     QAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNE
                     TLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSE
                     EWEAQRDSHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKA
                     SATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATDYLVQ
                     PFMDQLAFHQFYI"
BASE COUNT      359 a    310 c    333 g    390 t
ORIGIN
```

```
  Query Match           16.5%;  Score 248.2;  DB 9;  Length 1392;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        182 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 241

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        242 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 301

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        302 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 358

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        359 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 418

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        419 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 478

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db        479 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 538

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        539 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 598
```

```
Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        599 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 658

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        659 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 718

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        719 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 778

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db        779 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCCCAGGAGAGAAATGAAACGCTTTTTC 838

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db        839 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 885

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db        886 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 940

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db        941 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 998

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db        999 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1058

Qy        1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db        1059 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1108

Qy        1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db        1109 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1168

Qy        1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db        1169 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1228

Qy        1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db        1229 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1288

Qy        1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db        1289 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1348

Qy        1312 CCCCATTTGTTACACAAGT 1330
Db        1349 AGCCTTTTATGGACCAATT 1367
```

```
RESULT 13
AX195267
LOCUS       AX195267               1404 bp    DNA     linear   PAT 28-AUG-2001
DEFINITION  Sequence 3 from Patent WO0151671.
ACCESSION   AX195267
VERSION     AX195267.1  GI:15385819
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1404)
  AUTHORS   Mccarthy,J. and Cordell,B.
  TITLE     Methods for identifying inhibitors of neuronal degeneration
  JOURNAL   Patent: WO 0151671-A 3 19-JUL-2001;
            Scios Inc. (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .1404
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      362 a    312 c    337 g    393 t
ORIGIN
```

Query Match 16.5%; Score 248.2; DB 6; Length 1404;
Best Local Similarity 54.6%; Pred. No. 3e-46;
Matches 666; Conservative 0; Mismatches 513; Indels 40; Gaps 7;

```
Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        194 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 253

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        254 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 313

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        314 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 370

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        371 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 430

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        431 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 490

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db        491 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 550

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        551 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 610

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        611 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 670

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        671 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 730

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        731 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 790

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db        791 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 850

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db        851 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 897

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db        898 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 952

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db        953 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1010

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db        1011 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1070

Qy        1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
```

```
Db       1071 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1120

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138

Db       1121 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1180

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192

Db       1181 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1240

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252

Db       1241 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1300

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311

Db       1301 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1360

Qy       1312 CCCCATTTGTTACACAAGT 1330

Db       1361 AGCCTTTTATGGACCAATT 1379


RESULT 14
AR086330
LOCUS       AR086330               1488 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 3 from patent US 5985564.
ACCESSION   AR086330
VERSION     AR086330.1  GI:10013096
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1488)
  AUTHORS   Potter,H. and Li,J.
  TITLE     Assay for identifying agents that inhibit chromosome
            non-disjunction
  JOURNAL   Patent: US 5985564-A 3 16-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .1488
                     /organism="unknown"
BASE COUNT      374 a    321 c    354 g    439 t
ORIGIN

  Query Match          16.5%;  Score 248.2;  DB 6;  Length 1488;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178

Db         15 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 74

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238

Db         75 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 134

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298

Db        135 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 191

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358

Db        192 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 251

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418

Db        252 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 311

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478

Db        312 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 371

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538

Db        372 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 431

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598

Db        432 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 491

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658

Db        492 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 551

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718

Db        552 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 611

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778

Db        612 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 671

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838

Db        672 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 718

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898

Db        719 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 773

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958

Db        774 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 831

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018

Db        832 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 891

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078

Db        892 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 941

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138

Db        942 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1001

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192

Db       1002 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1061

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252

Db       1062 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1121

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311

Db       1122 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1181

Qy       1312 CCCCATTTGTTACACAAGT 1330

Db       1182 AGCCTTTTATGGACCAATT 1200


RESULT 15
AR095621
LOCUS       AR095621               2764 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 9 from patent US 6004794.
ACCESSION   AR095621
VERSION     AR095621.1  GI:10023656
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2764)
  AUTHORS   Karran,E.Howard., Clinkenbeard,H.Elizabeth., Browne,M.Joseph.,
            Southan,C.David., Creasy,C.Lee. and Livi,G.Pietro.
```

```
  TITLE     Human serine protease
  JOURNAL   Patent: US 6004794-A 9 21-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .2764
                     /organism="unknown"
BASE COUNT      715 a    623 c    652 g    773 t      1 others
ORIGIN

  Query Match           16.5%;  Score 248.2;  DB 6;  Length 2764;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||  | |  |  || ||   |    ||||||| || ||    || || ||    |
Db        442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||| | || |||||||   |||| ||   ||| || ||  |  | |  |
Db        502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
              ||||||  |  ||  |||   ||||  || | |     || ||        || || || |
Db        562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
              | | |  |||  |   ||  ||| ||  ||||||  | |||||  |  | ||  | |
Db        619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
              || | |||||  | || ||| | ||| | |||||||||| ||  ||||||  | || ||||
Db        679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
              |||||||||  |       | |  |  | || ||| | ||  |    || ||  ||   |
Db        739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
              ||||  | ||||  |   | ||  ||      |    |    |||    |      | | |
Db        799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
              | | ||| ||  | ||||||||  |  || || ||||||||||| | || || || |||
Db        859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
              || || ||||  ||   ||| || ||||| ||||| ||||||||||||| || ||||
Db        919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
              ||||||  |||| |  |  ||  ||| || || || |  ||| | || || ||  |    |
Db        979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
              | ||||| ||| |  |      |||| ||||| || |||||||||||  ||  |  |||| |
Db       1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
              | || || ||||| || ||   | |              | | | |||  || |  ||  |
Db       1099 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1145

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
              | || |      ||  ||| |       |  ||  || || | ||| |     |  | | |
Db       1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
              |    |     |   ||  | | | ||     ||  | | |||           |  |
Db       1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
                | |      |  | |     ||||        |   | | | ||  |     |
Db       1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
```

```
              |  | |||  |     |    ||| | |||           |     |  | ||
Db       1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
                || || || || || ||||||||  |||| || ||||||||||||  ||||||  | |
Db       1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
              | || || |      || |       |||||||||||  || || || ||||  || ||||
Db       1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
              |  | || ||| | ||| | ||  |  | || || ||| | |||||   |||| | || |
Db       1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
              ||||  |||| |||||   ||  || ||  |||| ||||||  ||    |   |  |
Db       1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy       1312 CCCCATTTGTTACACAAGT 1330
                || ||| |    ||| |
Db       1609 AGCCTTTTATGGACCAATT 1627


RESULT 16
AX244146
LOCUS       AX244146               2765 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 5 from Patent WO0167109.
ACCESSION   AX244146
VERSION     AX244146.1  GI:15859208
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2765)  .
  AUTHORS   Hale,R.S., Rowley,A.G. and Blackstock,W.G.
  TITLE     Method of identifying modulators of presenilin
  JOURNAL   Patent: WO 0167109-A 5 13-SEP-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .2765
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             249. .1652
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC88544.1"
                     /db_xref="GI:15859209"
                     /translation="MTELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRRSL
                     GHPEPLSNGRPQGNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                     VSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYK
                     VIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGP
                     LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                     ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNSKYNAESTERESQDTVAENDDG
                     GFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVL
                     VGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATD
                     YLVQPFMDQLAFHQFYI"
BASE COUNT      715 a    624 c    652 g    773 t      1 others
ORIGIN

  Query Match           16.5%;  Score 248.2;  DB 6;  Length 2765;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||  | |  |  || ||   |    ||||||| || ||    || || ||    |
Db        442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||| | || |||||||   |||| ||   ||| || ||  |  | |  |
Db        502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561
```

```
Qy       239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db       562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy       299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db       619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy       359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db       679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738

Qy       419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db       739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798

Qy       479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db       799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858

Qy       539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db       859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy       599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db       919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy       659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db       979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy       719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db      1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy       779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db      1099 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1145

Qy       839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db      1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy       899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db      1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy       959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db      1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy      1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db      1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy      1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db      1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy      1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db      1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy      1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db      1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy      1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db      1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy      1312 CCCCATTTGTTACACAAGT 1330
```

```
Db      1609 AGCCTTTTATGGACCAATT 1627
```

```
RESULT 17
AX244648
LOCUS       AX244648               2765 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 4 from Patent WO0167097.
ACCESSION   AX244648
VERSION     AX244648.1  GI:15859540
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2765)
  AUTHORS   Hale,R.S., Rowley,A.G. and Blackstock,W.G.
  TITLE     Novel assay
  JOURNAL   Patent: WO 0167097-A 4 13-SEP-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .2765
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             249. .1652
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC88578.1"
                     /db_xref="GI:15859541"
                     /translation="MTELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRRSL
                     GHPEPLSNGRPQGNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                     VSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYK
                     VIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGP
                     LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                     ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNSKYNAESTERESQDTVAENDDG
                     GFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVL
                     VGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATD
                     YLVQPFMDQLAFHQFYI"
BASE COUNT      715 a    624 c    652 g    773 t      1 others
ORIGIN

  Query Match           16.5%;  Score 248.2;  DB 6;  Length 2765;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy       119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db       442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy       179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db       502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561

Qy       239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db       562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy       299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db       619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy       359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db       679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738

Qy       419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db       739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798

Qy       479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db       799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858
```

```
Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db       1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1099 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1145

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1609 AGCCTTTTATGGACCAATT 1627
```

```
RESULT 18
HUMS182R
LOCUS       HUMS182R                2765 bp    mRNA    linear   PRI 04-OCT-1995
DEFINITION  Homo sapiens (clone cc33) S182 mRNA, complete cds.
ACCESSION   L42110
VERSION     L42110.1  GI:904118
KEYWORDS    .
SOURCE      Homo sapiens (clone: cc33) fetus brain cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2765)
  AUTHORS   Sherrington,R., Rogaev,E.I., Liang,Y., Rogaeva,E.A., Levesque,G.,
            Ikeda,M., Chi,H., Lin,C., Li,G., Holman,K., Tsuda,T., Mar,L.,
            Foncin,J.-F., Bruni,A.C., Montesi,M.P., Sorbi,S., Rainero,I.,
            Pinessi,L., Nee,L., Chumakov,I., Pollen,D., Brookes,A., Sanseau,P.,
            Polinsky,R.J., Wasco,W., Da Silva,H.A.R., Haines,J.L.,
            Pericak-Vance,M.A., Tanzi,R.E., Roses,A.D., Fraser,P.E.,
            Rommens,J.M. and St. George-Hyslop,P.H.
  TITLE     Cloning of a gene bearing missense mutations in early-onset
            familial Alzheimer's disease
  JOURNAL   Nature 375 (6534), 754-760 (1995)
  MEDLINE   95319502
   PUBMED   7596406
COMMENT     On Jul 25, 1995 this sequence version replaced gi:897613.
FEATURES             Location/Qualifiers
     source          1. .2765
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="14q24.3"
                     /clone="cc33"
                     /tissue_type="brain"
                     /dev_stage="fetus"
                     /germline
     gene            1. .2765
                     /gene="S182"
     mRNA            1. .2765
                     /gene="S182"
                     /note="map location near D14S77, D14S277, and D14S268"
     5'UTR           1. .248
                     /gene="S182"
     CDS             249. .1652
                     /gene="S182"
                     /codon_start=1
                     /protein_id="AAB46416.1"
                     /db_xref="GI:904119"
                     /translation="MTELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRRSL
                     GHPEPLSNGRPQGNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                     VSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYK
                     VIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGP
                     LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                     ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNSKYNAESTERESQDTVAENDDG
                     GFSEEWEAQRDSHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVL
                     VGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATD
                     YLVQPFMDQLAFHQFYI"
     3'UTR           1650. .2765
                     /gene="S182"
     polyA_site      2765
                     /gene="S182"
BASE COUNT      715 a    624 c    652 g    773 t      1 others
ORIGIN
```

```
  Query Match           16.5%;  Score 248.2;  DB 9;  Length 2765;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db        739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798
```

```
Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db       1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1099 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1145

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1609 AGCCTTTTATGGACCAATT 1627
```

```
RESULT 19
G27112
LOCUS       G27112                    2765 bp    DNA     linear   STS 15-JUN-1996
DEFINITION  human STS SHGC-31609, sequence tagged site.
ACCESSION   G27112
VERSION     G27112.1  GI:1375362
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2765)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu

            Primer A: CTAACAAAGTCAAGATTCCCGG
            Primer B: CAACTTCCGGGCCTATCATA
            STS size: 202
            PCR Profile:
                    Initial incubation: 94 degrees C for 90 seconds

                    Denaturation:       94 degrees C for 15 seconds
                    Annealing:          62 degrees C for 23 seconds
                    Polymerization:     72 degrees C for 30 seconds
                    PCR Cycles:         30
                    Thermal Cycler:     Perkin Elmer 9600
            Protocol:
                    Template:           25 ng
                    Primer:             each 1 uM
                    dNTPs:              each 200 uM
                    Taq Polymerase:     0.05 units/ul
                    Total Vol:          10 ul

            Buffer:
                    MgCl2:              2.5 mM
                    KCl:                50  mM
                    Tris-HCl:           20  mM
                    pH:                 8.3

            Prepared with primer pairs provided by Sandoz, derived from L42110
            -- Washington University/Merck EST sequence.
FEATURES                Location/Qualifiers
     source             1. .2765
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /map="14"
     STS                1741. .1942
     primer_bind        1741. .1762
     primer_bind        complement(1923. .1942)
BASE COUNT      715 a    624 c    652 g    773 t      1 others
ORIGIN
```

```
Query Match             16.5%;  Score 248.2;  DB 11;  Length 2765;
Best Local Similarity   54.6%;  Pred. No. 3e-46;
Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738
```

```
Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db        739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db       1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1099 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 1145

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1609 AGCCTTTTATGGACCAATT 1627
```

RESULT 20
AR060154
LOCUS AR060154 2791 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 133 from patent US 5840540.
ACCESSION AR060154
VERSION AR060154.1 GI:5986604
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2791)
AUTHORS St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
TITLE Nucleic acids encoding presenilin II
JOURNAL Patent: US 5840540-A 133 24-NOV-1998;
FEATURES Location/Qualifiers
source 1. .2791
/organism="unknown"
BASE COUNT 742 a 624 c 652 g 773 t
ORIGIN

Query Match 16.5%; Score 248.2; DB 6; Length 2791;
Best Local Similarity 54.6%; Pred. No. 3e-46;
Matches 666; Conservative 0; Mismatches 513; Indels 40; Gaps 7;

```
Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db        739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db       1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1099 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 1145

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
```

```
Db       1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1609 AGCCTTTTATGGACCAATT 1627
```

```
RESULT 21
AR087309
LOCUS       AR087309                2791 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 133 from patent US 5986054.
ACCESSION   AR087309
VERSION     AR087309.1  GI:10014072
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2791)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Genetic sequences and proteins related to alzheimer's disease
  JOURNAL   Patent: US 5986054-A 133 16-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .2791
                     /organism="unknown"
BASE COUNT      742 a    624 c    652 g    773 t
ORIGIN

  Query Match           16.5%;  Score 248.2;  DB 6;  Length 2791;
  Best Local Similarity  54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
```

```
Db        739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db       1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1099 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 1145

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1609 AGCCTTTTATGGACCAATT 1627
```

```
RESULT 22
AR134496
LOCUS       AR134496                2791 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 133 from patent US 6194153.
ACCESSION   AR134496
VERSION     AR134496.1  GI:14123401
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
```

```
          Unclassified.
REFERENCE   1  (bases 1 to 2791)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Methods for determining risk of developing alzheimer's disease by
            detecting mutations in the presenilin 1 (PS-1) gene
  JOURNAL   Patent: US 6194153-A 133 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .2791
                     /organism="unknown"
BASE COUNT      742 a    624 c    652 g    773 t
ORIGIN

  Query Match           16.5%;  Score 248.2;  DB 6;  Length 2791;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||  | |  |  || ||   |    ||||||| || ||    || || ||    |
Db        442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||| | || |||||||   |||| ||   ||| || ||  |  | | |
Db        502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
              ||||||  |  ||  |||   ||||  || | |      || ||        || || || |
Db        562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
                | |  |||  |   ||  ||| ||   ||||||  | |||||  |  | ||  | |
Db        619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
              || |  |||||  |  || ||| |  ||| |  |||||||||| ||  ||||||  | || ||||
Db        679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
                |||||||||  |      | |  |  | || ||| | ||  |    || ||  ||   |
Db        739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
              ||||  | ||||  |   | ||  ||     |    |   |||    |     |  |.|
Db        799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
              | | ||| ||  | |||||||| |  || || ||||||||||||| | || || || |||
Db        859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
                || || ||||  ||   ||| || ||||| ||||| ||||||||||||| || ||||
Db        919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
              ||||||  |||| |  |  ||  ||| || || || |  ||| | ||  || ||  |    |
Db        979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
              | ||||| ||| |  |      |||| ||||| || ||||||||||| ||  |  |||| |
Db       1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
              | || || ||||| || ||   | |            | | | |||  || |  ||  |
Db       1099 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1145

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
                | || |    ||  ||| |      |  ||  || || | ||| |    |  | | |
Db       1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
                |    |    |   ||  | | | ||    ||  | | |||          |  |
Db       1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
                | |    |  | | |    ||||       |   |  | | ||  |    |
Db       1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
              |   |  |||  |    |   ||| | |||                  |    |  | ||
Db       1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
                || || || || || ||||||||  |||| || ||||||||||||  ||||||  | |
Db       1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
              | || || |      || |       |||||||||||| || || || ||||  || ||||
Db       1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
              |  | || ||| | ||| | ||  |  | || || ||| | |||||   |||| | || |
Db       1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
              ||||  |||| |||||   ||  || ||  |||| ||||||   ||    |  |  |
Db       1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy       1312 CCCCATTTGTTACACAAGT 1330
                || ||| |    ||| |
Db       1609 AGCCTTTTATGGACCAATT 1627


RESULT 23
AR144597
LOCUS       AR144597               2792 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 133 from patent US 6210919.
ACCESSION   AR144597
VERSION     AR144597.1  GI:15106464
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2792)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Genetic sequences and proteins related to alzheimer's disease
  JOURNAL   Patent: US 6210919-A 133 03-APR-2001;
FEATURES             Location/Qualifiers
     source          1. .2792
                     /organism="unknown"
BASE COUNT      742 a    624 c    652 g    773 t      1 others
ORIGIN

  Query Match           16.5%;  Score 248.2;  DB 6;  Length 2792;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||  | |  |  || ||   |    ||||||| || ||    || || ||    |
Db        442 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 501

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||| | || |||||||   |||| ||   ||| || ||  |  | | |
Db        502 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 561

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
              ||||||  |  ||  |||   ||||  || | |      || ||        || || || |
Db        562 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 618

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
                | |  |||  |   ||  ||| ||   ||||||  | |||||  |  | ||  | |
Db        619 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 678

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
              || |  |||||  |  || ||| |  ||| |  |||||||||| ||  ||||||  | || ||||
Db        679 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 738

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
```

```
Db        739 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 798

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538

Db        799 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 858

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598

Db        859 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 918

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658

Db        919 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 978

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718

Db        979 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1038

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778

Db        1039 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1098

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838

Db        1099 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 1145

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898

Db        1146 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1200

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958

Db        1201 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1258

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018

Db        1259 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1318

Qy        1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078

Db        1319 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1368

Qy        1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138

Db        1369 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1428

Qy        1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192

Db        1429 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1488

Qy        1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252

Db        1489 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1548

Qy        1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311

Db        1549 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1608

Qy        1312 CCCCATTTGTTACACAAGT 1330

Db        1609 AGCCTTTTATGGACCAATT 1627


RESULT 24
BC011729
LOCUS       BC011729                2811 bp    mRNA    linear   PRI 02-AUG-2001
DEFINITION  Homo sapiens, clone MGC:19554 IMAGE:4423098, mRNA, complete cds.
ACCESSION   BC011729
VERSION     BC011729.1  GI:15079860
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2811)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site:       http://www.nisc.nih.gov/
            Contact:        nisc_mgc@nhgri.nih.gov
            Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
            Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 27 Row: p Column: 3
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 4506162.
FEATURES              Location/Qualifiers
     source           1: .2811
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="MGC:19554 IMAGE:4423098"
                      /tissue_type="Skin, melanotic melanoma."
                      /clone_lib="NIH_MGC_20"
                      /lab_host="DH10B-R"
                      /note="Vector: pOTB7"
     CDS              275. .1666
                      /codon_start=1
                      /product="Unknown (protein for MGC:19554)"
                      /protein_id="AAH11729.1"
                      /db_xref="GI:15079861"
                      /translation="MTELPAPLSYFQNAQMSEDNHLSNTNDNRERQEHNDRRSLGHPE
                      PLSNGRPQGNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFY
                      TRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHA
                      WLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQ
                      QAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNE
                      TLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSE
                      EWEAQRDSHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKA
                      SATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATDYLVQ
                      PFMDQLAFHQFYI"
BASE COUNT      749 a    626 c    662 g    774 t
ORIGIN

  Query Match           16.5%;  Score 248.2;  DB 9;  Length 2811;
  Best Local Similarity   54.6%;  Pred. No. 3e-46;
  Matches  666;  Conservative    0;  Mismatches  513;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178

Db        456 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 515

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238

Db        516 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 575

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
```

```
Db        576 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 632

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358

Db        633 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 692

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418

Db        693 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 752

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478

Db        753 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 812

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538

Db        813 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 872

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598

Db        873 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 932

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658

Db        933 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 992

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718

Db        993 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1052

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778

Db        1053 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1112

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838

Db        1113 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1159

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898

Db        1160 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1214

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958

Db        1215 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1272

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018

Db        1273 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1332

Qy        1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078

Db        1333 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1382

Qy        1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138

Db        1383 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1442

Qy        1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192

Db        1443 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1502

Qy        1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252

Db        1503 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1562

Qy        1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311

Db        1563 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1622

Qy        1312 CCCCATTTGTTACACAAGT 1330

Db        1623 AGCCTTTTATGGACCAATT 1641
```

```
RESULT 25
AX418176
LOCUS       AX418176                1404 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 1 from Patent WO0202601.
ACCESSION   AX418176
VERSION     AX418176.1  GI:21523186
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Carter,D.B. and Tomasselli,A.G.
  TITLE     Mutant presenilin 1 and presenilin 2 polypeptides
  JOURNAL   Patent: WO 0202601-A 1 10-JAN-2002;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1404
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       772. .777
                     /note="site directed mutagensis sites"
BASE COUNT      361 a    312 c    335 g    390 t      6 others
ORIGIN

  Query Match           16.5%;  Score 247;  DB 6;  Length 1404;
  Best Local Similarity   54.4%;  Pred. No. 5.7e-46;
  Matches  663;  Conservative    0;  Mismatches  516;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178

Db        194 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 253

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTTACGATGAACACGATTACGT 238

Db        254 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 313

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298

Db        314 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 370

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358

Db        371 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 430

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418

Db        431 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 490

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478

Db        491 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 550

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538

Db        551 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 610

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598

Db        611 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 670

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658

Db        671 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 730

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718

Db        731 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATNNNNNNGCTGTTTTGTGTC 790

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
```

```
Db        791 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 850
Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db        851 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 897
Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db        898 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 952
Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db        953 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1010
Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1011 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1070
Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1071 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1120
Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1121 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1180
Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1181 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1240
Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1241 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1300
Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1301 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1360
Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1361 AGCCTTTTATGGACCAATT 1379
```

```
RESULT 26
AR082573
LOCUS       AR082573                1914 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 24 from patent US 5973133.
ACCESSION   AR082573
VERSION     AR082573.1  GI:10009295
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1914).
  AUTHORS   Hardy,J.A. and Goate,A.M.
  TITLE     Mutant S182 genes
  JOURNAL   Patent: US 5973133-A 24 26-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .1914
                     /organism="unknown"
BASE COUNT       492 a    427 c    467 g    520 t      8 others
ORIGIN

  Query Match           16.4%;  Score 246.6;  DB 6;  Length 1914;
  Best Local Similarity  54.6%;  Pred. No. 7e-46;
  Matches  665;  Conservative    0;  Mismatches  514;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        401 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 460
Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        461 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 520
Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        521 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 577
Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        578 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 637
Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        638 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 697
Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db        698 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 757
Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        758 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 817
Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        818 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 877
Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        878 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 937
Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        938 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 997
Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db        998 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAGATGAAACGCTTTTTC 1057
Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1058 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1104
Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1105 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1159
Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1160 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1217
Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1218 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1277
Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1278 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGT----------ATCCTCGCTGGTGAAGACC 1327
Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1328 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1387
Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1388 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1447
Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1448 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1507
Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1508 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1567
```

```
Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1568 AGCCTTTTATGGACCAATT 1586


RESULT 27
AR060024
LOCUS       AR060024               2791 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 1 from patent US 5840540.
ACCESSION   AR060024
VERSION     AR060024.1  GI:5986474
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2791)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Nucleic acids encoding presenilin II
  JOURNAL   Patent: US 5840540-A 1 24-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .2791
                     /organism="unknown"
BASE COUNT      742 a    623 c    654 g    771 t      1 others
ORIGIN

  Query Match           16.4%;  Score 246.6;  DB 6;  Length 2791;
  Best Local Similarity   54.6%;  Pred. No. 7e-46;
  Matches  665;  Conservative    0;  Mismatches  514;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        443 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 502

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        503 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 562

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        563 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 619

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        620 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 679

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        680 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 739

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db        740 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 799

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        800 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 859

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        860 ATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 919

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        920 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 979

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        980 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1039

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db       1040 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1099

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1100 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1146

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1147 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1201

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1202 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1259

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1260 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1319

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1320 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1369

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1370 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1429

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1430 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1489

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1490 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1549

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1550 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1609

Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1610 AGCCTTTTATGGACCAATT 1628


RESULT 28
AR087179
LOCUS       AR087179               2791 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 1 from patent US 5986054.
ACCESSION   AR087179
VERSION     AR087179.1  GI:10013942
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2791)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Genetic sequences and proteins related to alzheimer's disease
  JOURNAL   Patent: US 5986054-A 1 16-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .2791
                     /organism="unknown"
BASE COUNT      742 a    623 c    654 g    771 t      1 others
ORIGIN

  Query Match           16.4%;  Score 246.6;  DB 6;  Length 2791;
  Best Local Similarity   54.6%;  Pred. No. 7e-46;
  Matches  665;  Conservative    0;  Mismatches  514;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        443 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 502

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        503 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 562
```

```
Qy       239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db       563 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 619

Qy       299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db       620 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 679

Qy       359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db       680 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 739

Qy       419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db       740 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 799

Qy       479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db       800 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 859

Qy       539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db       860 ATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 919

Qy       599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db       920 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 979

Qy       659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db       980 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1039

Qy       719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db      1040 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1099

Qy       779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db      1100 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 1146

Qy       839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db      1147 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1201

Qy       899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db      1202 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1259

Qy       959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db      1260 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1319

Qy      1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db      1320 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1369

Qy      1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db      1370 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1429

Qy      1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db      1430 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1489

Qy      1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db      1490 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1549

Qy      1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db      1550 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1609

Qy      1312 CCCCATTTGTTACACAAGT 1330
Db      1610 AGCCTTTTATGGACCAATT 1628
```

RESULT 29
AR134366

LOCUS	AR134366 2791 bp DNA linear PAT 16-MAY-2001
DEFINITION	Sequence 1 from patent US 6194153.
ACCESSION	AR134366
VERSION	AR134366.1 GI:14123271
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
	Unclassified.
REFERENCE	1 (bases 1 to 2791)
AUTHORS	St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
TITLE	Methods for determining risk of developing alzheimer's disease by detecting mutations in the presenilin 1 (PS-1) gene
JOURNAL	Patent: US 6194153-A 1 27-FEB-2001;
FEATURES	Location/Qualifiers
source	1. .2791
	/organism="unknown"
BASE COUNT	742 a 623 c 654 g 771 t 1 others
ORIGIN	

Query Match 16.4%; Score 246.6; DB 6; Length 2791;
Best Local Similarity 54.6%; Pred. No. 7e-46;
Matches 665; Conservative 0; Mismatches 514; Indels 40; Gaps 7;

```
Qy       119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db       443 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 502

Qy       179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db       503 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 562

Qy       239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db       563 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 619

Qy       299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db       620 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 679

Qy       359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db       680 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 739

Qy       419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db       740 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 799

Qy       479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db       800 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 859

Qy       539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db       860 ATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 919

Qy       599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db       920 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 979

Qy       659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db       980 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1039

Qy       719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db      1040 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1099
```

```
Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1100 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 1146

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1147 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1201

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1202 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1259

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1260 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1319

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1320 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1369

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1370 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1429

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1430 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1489

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1490 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1549

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1550 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1609

Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1610 AGCCTTTTATGGACCAATT 1628


RESULT 30
AR144467
LOCUS       AR144467               2791 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 1 from patent US 6210919.
ACCESSION   AR144467
VERSION     AR144467.1  GI:15106334
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2791)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Genetic sequences and proteins related to alzheimer's disease
  JOURNAL   Patent: US 6210919-A 1 03-APR-2001;
FEATURES             Location/Qualifiers
     source          1. .2791
                     /organism="unknown"
BASE COUNT      742 a    623 c    654 g    771 t      1 others
ORIGIN

  Query Match           16.4%;  Score 246.6;  DB 6;  Length 2791;
  Best Local Similarity   54.6%;  Pred. No. 7e-46;
  Matches  665;  Conservative    0;  Mismatches  514;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
Db        443 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 502

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
Db        503 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 562

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
Db        563 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 619

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
Db        620 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 679

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
Db        680 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 739

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
Db        740 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 799

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
Db        800 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 859

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
Db        860 ATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 919

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
Db        920 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 979

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
Db        980 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1039

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
Db       1040 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1099

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
Db       1100 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 1146

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
Db       1147 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1201

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
Db       1202 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1259

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
Db       1260 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1319

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
Db       1320 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1369

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
Db       1370 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1429

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
Db       1430 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1489

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1490 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1549

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1550 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1609
```

```
Qy       1312 CCCCATTTGTTACACAAGT 1330
              || ||| |    ||| |
Db       1610 AGCCTTTTATGGACCAATT 1628
```

RESULT 31
HUMPS1MRNA
LOCUS HUMPS1MRNA 3056 bp mRNA linear PRI 03-AUG-1996
DEFINITION Homo sapiens (clone cc44) senilin 1 (PS1; S182) mRNA, complete cds.
ACCESSION L76517
VERSION L76517.1 GI:1479973
KEYWORDS .
SOURCE Homo sapiens (clone: cc44) (tissue library: J.Rommens) adult colon tumor cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3056)
AUTHORS Sherrington,R., Rogaev,E.I., Liang,Y., Rogaeva,E.A., Levesque,G., Ikeda,M., Chi,H., Lin,C., Li,G., Holman,K. et al.
TITLE Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease
JOURNAL Nature 375 (6534), 754-760 (1995)
MEDLINE 95319502
PUBMED 7596406
REFERENCE 2 (bases 1 to 3056)
AUTHORS Rogaev,E.I., Sherrington,R., Rogaeva,E.I., Levesque,G., Liang,Y., Ikeda,M., Tsuda,T., Rogaeva,E., Chi,H., Holman,K., Rommens,J.M. and St George-Hyslop,P.H.
TITLE Genomic structure, promoter sequence, and alternative splicing of the presenilin-1 gene associated with early onset Alzheimer's Disease
JOURNAL Unpublished (1996)
FEATURES Location/Qualifiers
source 1. .3056
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="14q24.3"
/clone="cc44"
/cell_line="CaCo2"
/cell_type="epithelial cell"
/tissue_type="colon tumor"
/dev_stage="adult"
/germline
/tissue_lib="J.Rommens"
gene 1. .3056
/gene="PS1"
CDS 554. .1945
/gene="PS1"
/function="unknown"
/note="Note alternate 5' start exon, and lacks Val26, Arg27, Ser28, and Gln29 in clone cc33 due to alternative splicing of exon 4.; putative"
/citation=[2]
/codon_start=1
/product="presenilin 1"
/protein_id="AAB46370.1"
/db_xref="GI:1479974"
/translation="MTELPAPLSYFQNAQMSEDNHLSNTNDNRERQEHNDRRSLGHPE
PLSNGRPQGNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFY
TRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHA
WLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNLGVVGMISIHWKGPLRLQ
QAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNE
TLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNSKYNAESTERESQDTVAENDDGGFSE
EWEAQRDSHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKA
SATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATDYLVQ
PFMDQLAFHQFYI"
polyA_site 3056
/gene="PS1"
BASE COUNT 762 a 688 c 740 g 866 t
ORIGIN

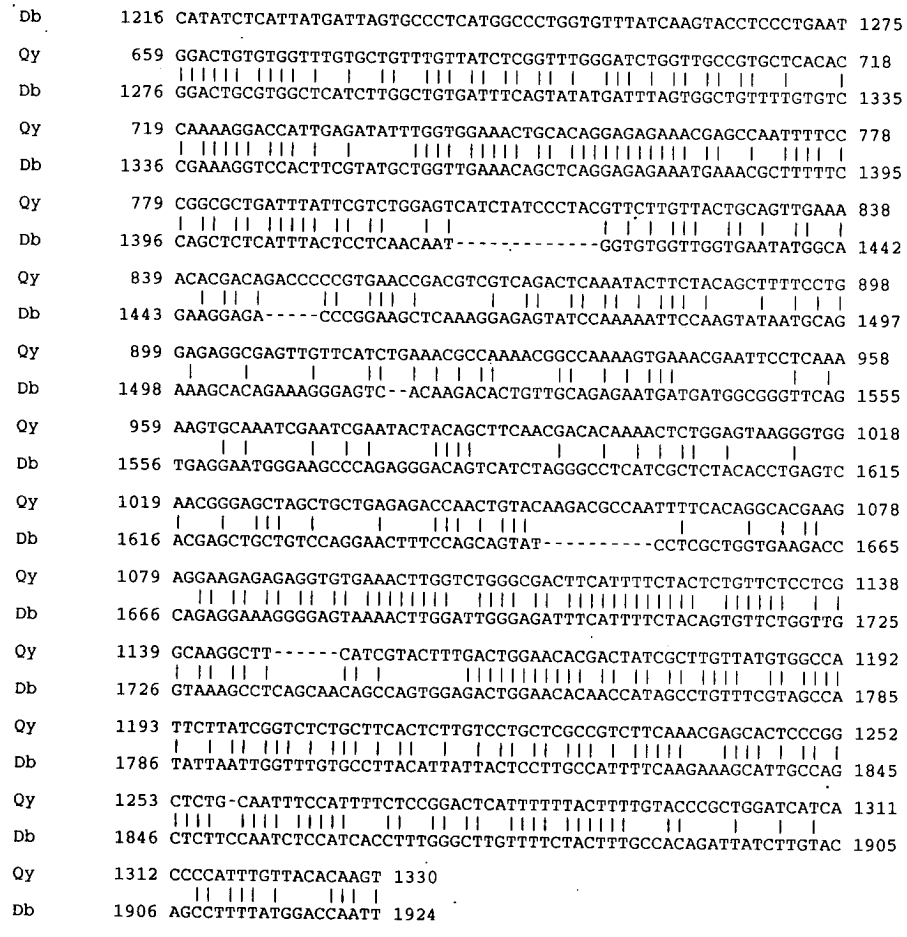
Query Match 16.4%; Score 246.6; DB 9; Length 3056;

Best Local Similarity 54.6%; Pred. No. 7e-46;
Matches 665; Conservative 0; Mismatches 514; Indels 40; Gaps 7;

```
Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||  | |  |  || ||   |    ||||||| || ||    || || ||    |
Db        735 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 794

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||| | || |||||||    |||| ||   ||| || ||  |  | |  |
Db*       795 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 854

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
              ||||||  |  ||  |||  |||| || | |     || ||           || || || |
Db        855 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 911

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
              | | |  |||  |  ||  ||| ||  ||||||  | ||||| |  |  ||  | |
Db        912 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 971

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
              || | |||||  | || ||| | ||| | ||||||||||  ||  ||||||  | || ||||
Db        972 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 1031

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
                |||||||||  |      | |  |  | || ||| | ||  |    ||  ||  ||  |
Db        1032 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 1091

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
              ||||  | ||||  |   | ||  ||      |    |||    |     |  |  |
Db        1092 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 1151

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
              | |  || ||  | |||||||| |  || || ||||||||||||| | || || || |||
Db        1152 ATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 1211

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
                || || ||||  ||   ||| || ||||| ||||| |||||||||||||||| || ||||
Db        1212 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 1271

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
              ||||||  |||| |  |  ||  ||| || || || |  ||| | || || ||  |  |
Db        1272 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1331

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
              | ||||| ||| |  |      |||| ||||| || |||||||||||  ||  |  |||| |
Db        1332 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1391

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
              | || || ||||| || ||   | |                | | | |||  ||  |
Db        1392 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1438

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
              | || |    ||  ||| |      |  ||  || || | ||| |    |  | | |
Db        1439 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1493

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
              |   |      |    ||  | | | ||     ||  | | |||           | |
Db        1494 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1551

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
              | |      |  | |     ||||        |   |  | | ||  |    |
Db        1552 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1611

Qy        1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
               |   | |||  |     | ||| | |||                |     | | ||
Db        1612 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1661

Qy        1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
               || || || || || |||||||| |||| || ||||||||||| ||||||  | |
Db        1662 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1721

Qy        1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
               | || || |      || |       |||||||||||| || || || ||||  || ||||
```

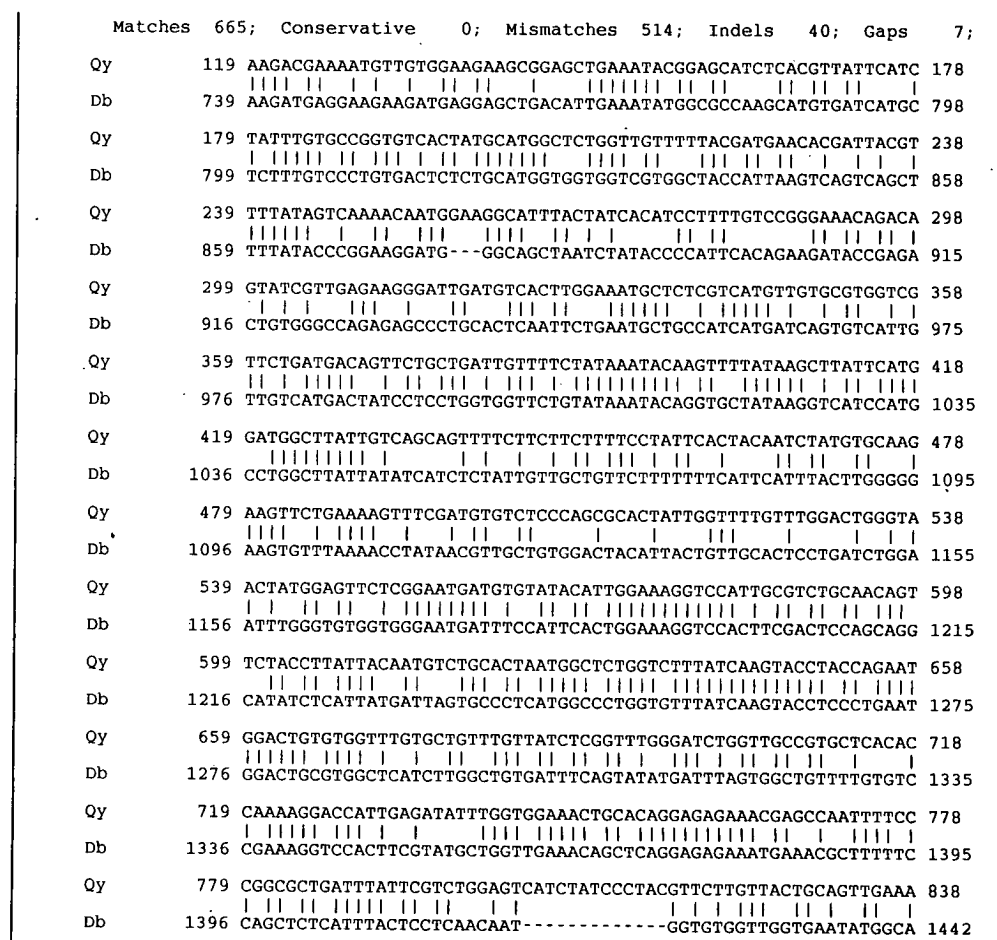
```
Db       1722 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1781

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252

Db       1782 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1841

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311

Db       1842 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1901

Qy       1312 CCCCATTTGTTACACAAGT 1330

Db       1902 AGCCTTTTATGGACCAATT 1920


RESULT 32
AR060026
LOCUS       AR060026               3087 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 5 from patent US 5840540.
ACCESSION   AR060026
VERSION     AR060026.1  GI:5986476
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3087)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Nucleic acids encoding presenilin II
  JOURNAL   Patent: US 5840540-A 5 24-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .3087
                     /organism="unknown"
BASE COUNT      789 a    688 c    740 g    866 t      4 others
ORIGIN

  Query Match           16.4%;  Score 246.6;  DB 6;  Length 3087;
  Best Local Similarity   54.6%;  Pred. No. 7e-46;
  Matches  665;  Conservative    0;  Mismatches  514;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178

Db        739 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 798

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238

Db        799 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 858

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298

Db        859 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 915

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358

Db        916 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 975

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418

Db        976 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 1035

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478

Db       1036 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 1095

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538

Db       1096 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 1155

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598

Db       1156 ATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 1215

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658

Db       1216 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 1275

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718

Db       1276 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1335

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778

Db       1336 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1395

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838

Db       1396 CAGCTCTCATTTACTCCTCAACAAT----------GGTGTGGTTGGTGAATATGGCA 1442

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898

Db       1443 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1497

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958

Db       1498 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1555

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018

Db       1556 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1615

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078

Db       1616 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1665

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138

Db       1666 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1725

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192

Db       1726 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1785

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252

Db       1786 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1845

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311

Db       1846 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1905

Qy       1312 CCCCATTTGTTACACAAGT 1330

Db       1906 AGCCTTTTATGGACCAATT 1924


RESULT 33
AR087181
LOCUS       AR087181               3087 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 5 from patent US 5986054.
ACCESSION   AR087181
VERSION     AR087181.1  GI:10013944
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3087)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Genetic sequences and proteins related to alzheimer's disease
  JOURNAL   Patent: US 5986054-A 5 16-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .3087
                     /organism="unknown"
BASE COUNT      789 a    688 c    740 g    866 t      4 others
ORIGIN

  Query Match           16.4%;  Score 246.6;  DB 6;  Length 3087;
  Best Local Similarity   54.6%;  Pred. No. 7e-46;
```

```
  Matches  665;  Conservative    0;  Mismatches  514;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||  | |  |  || ||   |    ||||||| || ||    || || ||    |
Db        739 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 798

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||| | || |||||||   |||| ||   ||| || ||  |  | |  |
Db        799 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 858

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
              ||||||  |  ||  |||   |||| || | |   || ||        || || || |
Db        859 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 915

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
              | | |   |||  |   ||   ||| ||   |||||| | ||||| |  | ||  | |
Db        916 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 975

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
              || | |||||  | || ||| | ||| | ||||||||||  ||  |||||| | || ||||
Db        976 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 1035

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
              |||||||||  |      | |  |  | || ||| | ||  |    || ||  ||   |
Db       1036 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 1095

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
              ||||  | ||||  |  | ||  ||      |   |||   |    |  | |  |  | |
Db       1096 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 1155

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
              | |  || ||  | |||||||| |  || || ||||||||||| | || || || |||
Db       1156 ATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 1215

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
              || || ||||  ||   ||| || ||||| ||||| |||||||||||||| || ||||
Db       1216 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 1275

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
              |||||| |||| |  |  ||  ||| || || || |  ||| | || || ||  |    |
Db       1276 GGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGTC 1335

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
              | ||||| ||| |  |    |||| ||||| || |||||||||| ||  |  |||| |
Db       1336 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 1395

Qy        779 CGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTGAAA 838
              | || || ||||| || ||   | |             | | | ||| || |  ||  |
Db       1396 CAGCTCTCATTTACTCCTCAACAAT------------GGTGTGGTTGGTGAATATGGCA 1442

Qy        839 ACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTCCTG 898
              | || |     ||  ||| |       |  ||  || || | ||| |    |  | | |
Db       1443 GAAGGAGA-----CCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAG 1497

Qy        899 GAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCTCAAA 958
              |   |    |   ||  | | | ||     ||  | |  |||          |  |
Db       1498 AAAGCACAGAAAGGGAGTC--ACAAGACACTGTTGCAGAGAATGATGATGGCGGGTTCAG 1555

Qy        959 AAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGGGTGG 1018
              | |     |  | |     ||||       |   |  | | ||  |     |
Db       1556 TGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATCGCTCTACACCTGAGTC 1615

Qy       1019 AACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCACGAAG 1078
              |   | |||  |    |  ||| | |||                |   | | ||
Db       1616 ACGAGCTGCTGTCCAGGAACTTTCCAGCAGTAT----------CCTCGCTGGTGAAGACC 1665

Qy       1079 AGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTCCTCG 1138
                || || || || || |||||||| |||| || ||||||||||  ||||||  | |
Db       1666 CAGAGGAAAGGGGAGTAAAACTTGGATTGGGAGATTTCATTTTCTACAGTGTTCTGGTTG 1725

Qy       1139 GCAAGGCTT------CATCGTACTTTGACTGGAACACGACTATCGCTTGTTATGTGGCCA 1192
              | || || |      || |       |||||||||||| || || || ||||  || ||||
Db       1726 GTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCA 1785

Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
              |  | || ||| | ||| | ||  |  | || || ||| | |||||  |||| | || |
Db       1786 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1845

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
              ||||  |||| |||||   ||  || ||  |||| ||||||   ||    |  |  |
Db       1846 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1905

Qy       1312 CCCCATTTGTTACACAAGT 1330
                || ||| |    ||| |
Db       1906 AGCCTTTTATGGACCAATT 1924


RESULT 34
AR134368
LOCUS       AR134368               3087 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 5 from patent US 6194153.
ACCESSION   AR134368
VERSION     AR134368.1  GI:14123273
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3087)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Methods for determining risk of developing alzheimer's disease by
            detecting mutations in the presenilin 1 (PS-1) gene
  JOURNAL   Patent: US 6194153-A 5 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .3087
                     /organism="unknown"
BASE COUNT      789 a    688 c    740 g    866 t      4 others
ORIGIN

  Query Match           16.4%;  Score 246.6;  DB 6;  Length 3087;
  Best Local Similarity   54.6%;  Pred. No. 7e-46;
  Matches  665;  Conservative    0;  Mismatches  514;  Indels   40;  Gaps    7;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||  | |  |  || ||   |    ||||||| || ||    || || ||    |
Db        739 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 798

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||| | || |||||||   |||| ||   ||| || ||  |  | |  |
Db        799 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 858

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
              ||||||  |  ||  |||   |||| || | |   || ||        || || || |
Db        859 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 915

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
              | | |   |||  |   ||   ||| ||   |||||| | ||||| |  | ||  | |
Db        916 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 975

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
              || | |||||  | || ||| | ||| | ||||||||||  ||  |||||| | || ||||
Db        976 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 1035

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
              |||||||||  |      | |  |  | || ||| | ||  |    || ||  ||   |
Db       1036 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 1095

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
              ||||  | ||||  |  | ||  ||      |   |||   |    |  | |  |  | |
Db       1096 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 1155

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
              | |  || ||  | |||||||| |  || || ||||||||||| | || || || |||
Db       1156 ATTTGGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 1215

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
              || || ||||  ||   ||| || ||||| ||||| |||||||||||||| || ||||
```



RESULT 35
AR144469
LOCUS AR144469 3087 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6210919.
ACCESSION AR144469
VERSION AR144469.1 GI:15106336
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 3087)
AUTHORS St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
TITLE Genetic sequences and proteins related to alzheimer's disease
JOURNAL Patent: US 6210919-A 5 03-APR-2001;
FEATURES Location/Qualifiers
source 1. .3087
/organism="unknown"
BASE COUNT 789 a 688 c 740 g 866 t 4 others
ORIGIN

Query Match 16.4%; Score 246.6; DB 6; Length 3087;
Best Local Similarity 54.6%; Pred. No. 7e-46;



Matches 665; Conservative 0; Mismatches 514; Indels 40; Gaps 7;

```
Qy       1193 TTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTCCCGG 1252
Db       1786 TATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAG 1845

Qy       1253 CTCTG-CAATTTCCATTTTCTCCGGACTCATTTTTTACTTTTGTACCCGCTGGATCATCA 1311
Db       1846 CTCTTCCAATCTCCATCACCTTTGGGCTTGTTTTCTACTTTGCCACAGATTATCTTGTAC 1905

Qy       1312 CCCCATTTGTTACACAAGT 1330
Db       1906 AGCCTTTTATGGACCAATT 1924


RESULT 36
AR060155
LOCUS       AR060155               1962 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 135 from patent US 5840540.
ACCESSION   AR060155
VERSION     AR060155.1  GI:5986605
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1962)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Nucleic acids encoding presenilin II
  JOURNAL   Patent: US 5840540-A 135 24-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .1962
                     /organism="unknown"
BASE COUNT      503 a    503 c    496 g    460 t
ORIGIN

  Query Match           16.2%;  Score 243.2;  DB 6;  Length 1962;
  Best Local Similarity  54.4%;  Pred. No. 4.2e-45;
  Matches  644;  Conservative    0;  Mismatches  498;  Indels   42;  Gaps    6;

Qy        116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
Db        376 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 435

Qy        176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTA 235
Db        436 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 495

Qy        236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
Db        496 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 552

Qy        296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
Db        553 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 612

Qy        356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
Db        613 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 672

Qy        416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
Db        673 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 732

Qy        476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
Db        733 GGGAAGTATTTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATCT 792

Qy        536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
Db        793 GGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 852

Qy        596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
Db        853 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 912

Qy        656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
Db        913 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 972

Qy        716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
Db        973 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1032

Qy        776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
Db       1033 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1089

Qy        836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
Db       1090 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1149

Qy        896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
Db       1150 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1209

Qy        955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
Db       1210 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1245

Qy       1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
Db       1246 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1301

Qy       1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
Db       1302 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1361

Qy       1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
Db       1362 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTA 1421

Qy       1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
Db       1422 GCCATACTGATCGGCCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTG 1481

Qy       1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
Db       1482 CCAGCCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTT 1525


RESULT 37
AR087310
LOCUS       AR087310               1962 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 135 from patent US 5986054.
ACCESSION   AR087310
VERSION     AR087310.1  GI:10014073
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1962)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Genetic sequences and proteins related to alzheimer's disease
  JOURNAL   Patent: US 5986054-A 135 16-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .1962
                     /organism="unknown"
BASE COUNT      503 a    503 c    496 g    460 t
ORIGIN

  Query Match           16.2%;  Score 243.2;  DB 6;  Length 1962;
  Best Local Similarity  54.4%;  Pred. No. 4.2e-45;
  Matches  644;  Conservative    0;  Mismatches  498;  Indels   42;  Gaps    6;

Qy        116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
Db        376 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 435
```

```
Qy       176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTTACGATGAACACGATTA 235
Db       436 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 495

Qy       236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
Db       496 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 552

Qy       296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
Db       553 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 612

Qy       356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
Db       613 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 672

Qy       416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
Db       673 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 732

Qy       476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
Db       733 GGGAAGTATTTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATCT 792

Qy       536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
Db       793 GGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 852

Qy       596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
Db       853 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 912

Qy       656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
Db       913 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 972

Qy       716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
Db       973 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1032

Qy       776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
Db      1033 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1089

Qy       836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
Db      1090 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1149

Qy       896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
Db      1150 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1209

Qy       955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
Db      1210 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1245

Qy      1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
Db      1246 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1301

Qy      1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
Db      1302 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1361

Qy      1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
Db      1362 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTA 1421

Qy      1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
Db      1422 GCCATACTGATCGGCCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTG 1481
```

```
Qy      1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
Db      1482 CCAGCCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTT 1525


RESULT 38
AR134497
LOCUS       AR134497               1962 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 135 from patent US 6194153.
ACCESSION   AR134497
VERSION     AR134497.1  GI:14123402
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1962)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Methods for determining risk of developing alzheimer's disease by
            detecting mutations in the presenilin 1 (PS-1) gene
  JOURNAL   Patent: US 6194153-A 135 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .1962
                     /organism="unknown"
BASE COUNT      503 a    503 c    496 g    460 t
ORIGIN

  Query Match           16.2%;  Score 243.2;  DB 6;  Length 1962;
  Best Local Similarity   54.4%;  Pred. No. 4.2e-45;
  Matches  644;  Conservative    0;  Mismatches  498;  Indels   42;  Gaps    6;

Qy       116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
Db       376 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 435

Qy       176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTTACGATGAACACGATTA 235
Db       436 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 495

Qy       236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
Db       496 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 552

Qy       296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
Db       553 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 612

Qy       356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
Db       613 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 672

Qy       416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
Db       673 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 732

Qy       476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
Db       733 GGGAAGTATTTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATCT 792

Qy       536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
Db       793 GGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 852

Qy       596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
Db       853 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 912

Qy       656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
Db       913 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 972

Qy       716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
Db       973 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1032
```

```
Qy        776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
              | || || || || ||||| ||   | | | |     |  ||   | |  ||| || |
Db       1033 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1089

Qy        836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
              |      |  |     | | ||| | |  |  | | |  || ||    | |||
Db       1090 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1149

Qy        896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
                |      ||  ||| | |||| |   | | |  |||||    |
Db       1150 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1209

Qy        955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
              |  |||| |  | ||  | |      |  ||| ||
Db       1210 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1245

Qy        1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
               |  |||  |  ||||     |      |    || | |    ||  |
Db       1246 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1301

Qy        1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
               ||   ||| || ||||| || |||||||| ||||| || ||||||||||||  ||||||
Db       1302 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1361

Qy        1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
               | || ||||| ||| |   |          |||||||||| || || || || | |||
Db       1362 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTA 1421

Qy        1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
               ||||| || ||||| || ||| | ||  |  ||||||||||| | |||||   |||  |
Db       1422 GCCATACTGATCGGCCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTG 1481

Qy        1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
               || || ||  | || |||||   || ||| ||| | || |||||
Db       1482 CCAGCCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTT 1525
```

RESULT 39
AR144598
LOCUS AR144598 1964 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 135 from patent US 6210919.
ACCESSION AR144598
VERSION AR144598.1 GI:15106465
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1964)
AUTHORS St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
TITLE Genetic sequences and proteins related to alzheimer's disease
JOURNAL Patent: US 6210919-A 135 03-APR-2001;
FEATURES Location/Qualifiers
source 1. .1964
/organism="unknown"
BASE COUNT 503 a 503 c 496 g 460 t 2 others
ORIGIN

Query Match 16.2%; Score 243.2; DB 6; Length 1964;
Best Local Similarity 54.4%; Pred. No. 4.2e-45;
Matches 644; Conservative 0; Mismatches 498; Indels 42; Gaps 6;

```
Qy         116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
               || |||| ||  | |  |  |||||   |    ||||||| |||||    || || ||
Db         378 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 437

Qy         176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTA 235
               || ||||| || ||| | || |||||||   | || ||    || || ||  |  | |
Db         438 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 497

Qy         236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
               || ||||  |  ||  | |   | ||  || | |     || ||       || || |
Db         498 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 554

Qy         296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
               | | | | |   | |  |  ||  ||  |    |||||   | ||||| |  | ||
Db         555 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 614

Qy         356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
               | ||  | |||||  | || ||| | ||  | ||||||||||| ||  || ||| | || |
Db         615 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 674

Qy         416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
               | |  |||||||||  |     | |  |  | || ||| | ||  |    || ||  |
Db         675 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 734

Qy         476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
               |||||  | || |  | | |||| |       |    |    ||      |  | |
Db         735 GGGAAGTATTTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATCT 794

Qy         536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
               | || | ||| ||  |||| |||||    || || |||||||| ||  | || ||||| |
Db         795 GGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 854

Qy         596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
               ||   || || |||  ||   ||| || ||||| ||||| ||||||||||||||| || |
Db         855 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 914

Qy         656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
               ||||||| |  ||| |  |  || ||| || || || |  ||| |||| || || ||  |
Db         915 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 974

Qy         716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
               || ||||| ||| |  |     |||| ||||| || ||||| ||||| ||| |  | |
Db         975 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1034

Qy         776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
               | || || || || ||||| ||   | | | |     |  ||   | |  ||| || |
Db        1035 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1091

Qy         836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
               |      |  |     | | ||| | |  |  | | |  || ||    | |||
Db        1092 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1151

Qy         896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
                |      ||  ||| | |||| |   | | |  |||||    |
Db        1152 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1211

Qy         955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
               |  |||| |  | ||  | |      |  ||| ||
Db        1212 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1247

Qy         1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
                |  |||  |  ||||     |      |    || | |    ||  |
Db        1248 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1303

Qy         1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
                ||   ||| || ||||| || |||||||| ||||| || ||||||||||||  ||||||
Db        1304 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1363

Qy         1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
                | || ||||| ||| |   |          |||||||||| || || || || | |||
Db        1364 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTA 1423

Qy         1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
                ||||| || ||||| || ||| | ||  |  ||||||||||| | |||||   |||  |
Db        1424 GCCATACTGATCGGCCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTG 1483

Qy         1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
                || || ||  | || |||||   || ||| ||| | || |||||
Db        1484 CCAGCCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTT 1527
```

RESULT 40
MUSS1PR
LOCUS MUSS1PR 1964 bp mRNA linear ROD 24-JUL-1995

```
DEFINITION  Mus musculus S182 protein mRNA, complete cds.
ACCESSION   L42177
VERSION     L42177.1  GI:904129
KEYWORDS    S182 protein.
SOURCE      Mus musculus (clone: 917-935) adult brain cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1964)
  AUTHORS   Sherrington,R., Rogaev,E.I., Liang,Y., Rogaeva,E.A., Levesque,G.,
            Ikeda,M., Chi,H., Lin,C., Li,G., Holman,K., Tsuda,T., Mar,L.,
            Foncin,J.-F., Bruni,A.C., Montesi,M.P., Sorbl,S., Rainero,I.,
            Pinessi,L., Nee,L., Chumakov,I., Pollen,D., Brookes,A., Sanseau,P.,
            Polinsky,R.J., Wasco,W., Da Silva,H.A.R., Haines,J.L.,
            Pericak-Vance,M.A., Tanzi,R.E., Roses,A.D., Fraser,P.E.,
            Rommens,J.M. and St. George-Hyslop,P.H.
  TITLE     Cloning of a gene bearing missense mutations in early-onset
            familial Alzheimer's disease
  JOURNAL   Nature 375 (6534), 754-760 (1995)
  MEDLINE   95319502
   PUBMED   7596406
COMMENT     On Jul 25, 1995 this sequence version replaced gi:897616.
FEATURES             Location/Qualifiers
     source          1. .1964
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="14q24.3"
                     /clone="917-935"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /germline
                     /note="clone:  composite sequence from a ~ 0.7 kb RT-PCR
                     clone and an overlapping ~2.0 kb cDNA clone mc6-3-2
                     recovered from adult mouse brain cDNA library in
                     Bluescript at the XhoI site."
     mRNA            1. .1964
     5'UTR           1. .187
     CDS             188. .1591
                     /function="unknown"
                     /note="murine homologue of human S182 gene (accession
                     number L42110); putative"
                     /codon_start=1
                     /product="S182 protein"
                     /protein_id="AAC42094.1"
                     /db_xref="GI:904130"
                     /translation="MTEIPAPLSYFQNAQMSEDSHSSSAIRSQNDSQERQQQHDRQRL
                     DNPEPISNGRPQSNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                     VSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVIMTILLVVLYKYRCYK
                     VIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYVTVALLIWNFGVVGMIAIHWKGP
                     LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                     ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVPKNPKYNTQRAERETQDSGSGNDDG
                     GFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSILTSEDPEERGVKLGLGDFIFYSVL
                     VGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATD
                     YLVQPFMDQLAFHQFYI"
     3'UTR           1592. .1964
BASE COUNT      503 a    503 c    496 g    460 t      2 others
ORIGIN
```

```
  Query Match           16.2%;  Score 243.2;  DB 10;  Length 1964;
  Best Local Similarity   54.4%;  Pred. No. 4.2e-45;
  Matches  644;  Conservative    0;  Mismatches  498;  Indels   42;  Gaps    6;

Qy        116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
              || |||| ||  | |  |  |||||    |    ||||||| |||||    || || ||
Db        378 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 437

Qy        176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTA 235
                || ||||| || ||| | || |||||||   | || ||    || || ||  |  | |
Db        438 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 497

Qy        236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
                || ||||  |  ||  | |   | ||  || | |     || ||          || || |
Db        498 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 554

Qy        296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
              | | | | |   | |  |   ||   ||  |    |||||   | ||||| | | ||
Db        555 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 614

Qy        356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
              | ||  | |||||  | || ||| | ||  | ||||||||| ||  || ||| | || |
Db        615 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 674

Qy        416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
              | |  |||||||||| |       | | |  | || ||| | ||  |    || ||  |
Db        675 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 734

Qy        476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
                |||||  | || |  | | ||||  |      |   |   ||     |   | |
Db        735 GGGAAGTATTTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATCT 794

Qy        536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
              | || | ||| ||  |||| |||||    || || |||||||| ||  | || |||||
Db        795 GGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 854

Qy        596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
              ||   || || |||| ||   ||| || ||||| ||||| ||||||||||||||| || |
Db        855 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 914

Qy        656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
              |||||||  |  ||| |  |  ||  ||| || || || |  ||| |||| || ||  |
Db        915 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 974

Qy        716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
              || ||||| ||| |  |      |||| ||||| || ||||| ||||| ||| |  | |
Db        975 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1034

Qy        776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
              | || || || || ||||| ||   | |  | |     |  ||   | |  ||| || |
Db       1035 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1091

Qy        836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
              |       |  |      | | ||| | |    |  | |  || ||    |.|||.
Db       1092 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1151

Qy        896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
              |       |   ||  ||| | ||||   |   | |  |||||   |
Db       1152 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1211

Qy        955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
              |  |||| |  | ||   | |     |  ||| ||
Db       1212 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1247

Qy        1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
                   | |||  |  ||||     |      |    |    || | |    ||  |
Db       1248 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1303

Qy        1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
               ||   ||| || ||||| || |||||||| ||||| || ||||||||||||  ||||||
Db       1304 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1363

Qy        1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
                 | || ||||| ||| |   |           ||||||||||| || || || || | |||
Db       1364 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTA 1423

Qy        1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
               ||||| || ||||| || ||| | ||  |  ||||||||||| | |||||   ||| |
Db       1424 GCCATACTGATCGGCCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTG 1483

Qy        1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
               || || ||  | || |||||   || ||| ||| | || ||||
Db       1484 CCAGCCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTT 1527
```

```
RESULT 41
BC030409
LOCUS       BC030409                2016 bp    mRNA    linear   ROD 07-AUG-2002
```

```
DEFINITION  Mus musculus, presenilin 1, clone MGC:40637 IMAGE:5133302, mRNA,
            complete cds.
ACCESSION   BC030409
VERSION     BC030409.1  GI:20987825
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2016)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 65 Row: d Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6679492.
FEATURES              Location/Qualifiers
     source           1. .2016
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /map="FVB/N"
                      /clone="MGC:40637 IMAGE:5133302"
                      /tissue_type="Liver, normal. 5 month old male mouse."
                      /clone_lib="NCI_CGAP_Li9"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     CDS              167. .1570
                      /codon_start=1
                      /product="presenilin 1"
                      /protein_id="AAH30409.1"
                      /db_xref="GI:20987826"
                      /db_xref="LocusID:19164"
                      /translation="MTEIPAPLSYFQNAQMSEDSHSSSAIRSQNDSQERQQQHDRQRL
                      DNPEPISNGRPQSNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                      VSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVIMTILLVVLYKYRCYK
                      VIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYVTVALLIWNFGVVGMIAIHWKGP
                      LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                      ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVPKNPKYNTQRAERETQDSGSGNDDG
                      GFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSILTSEDPEERGVKLGLGDFIFYSVL
                      VGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATD
                      YLVQPFMDQLAFHQFYI"
BASE COUNT      531 a    516 c    498 g    471 t
ORIGIN

  Query Match          16.2%;  Score 243.2;  DB 10;  Length 2016;
  Best Local Similarity  54.4%;  Pred. No. 4.2e-45;
  Matches  644;  Conservative    0;  Mismatches  498;  Indels   42;  Gaps    6;

Qy       116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
Db       357 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 416

Qy       176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTA 235
Db       417 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 476

Qy       236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
Db       477 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 533

Qy       296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
Db       534 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 593

Qy       356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
Db       594 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 653

Qy       416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
Db       654 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 713

Qy       476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
Db       714 GGGAAGTATTTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATCT 773

Qy       536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
Db       774 GGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 833

Qy       596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
Db       834 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 893

Qy       656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
Db       894 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 953

Qy       716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
Db       954 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1013

Qy       776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
Db       1014 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1070

Qy       836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
Db       1071 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1130

Qy       896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
Db       1131 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1190

Qy       955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
Db       1191 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1226

Qy       1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
Db       1227 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1282

Qy       1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
Db       1283 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1342

Qy       1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
Db       1343 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTA 1402

Qy       1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
```

```
Db       1403 GCCATACTGATCGGCCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTG 1462

Qy       1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
Db       1463 CCAGCCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTT 1506


RESULT 42
A63557
LOCUS       A63557                  2681 bp    DNA     linear   PAT 12-MAR-1998
DEFINITION  Sequence 10 from Patent WO9722695.
ACCESSION   A63557
VERSION     A63557.1  GI:3717212
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2681)
  AUTHORS   Talerman,A., Amson,R. and Cohen,D.
  TITLE     NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC AGENTS FOR
            TREATING CANCER
  JOURNAL   Patent: WO 9722695-A 10 26-JUN-1997;
            FONDATION JEAN DAUSSET CEPH (FR)
COMMENT     Other publication FR 2747691 19971024
            Other publication FR 2742766 19970627.
FEATURES             Location/Qualifiers
     source          1. .2681
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      663 a    709 c    658 g    650 t      1 others
ORIGIN

  Query Match           16.2%;  Score 243.2;  DB 6;  Length 2681;
  Best Local Similarity  54.4%;  Pred. No. 4.2e-45;
  Matches  644;  Conservative    0;  Mismatches  498;  Indels   42;  Gaps    6;

Qy        116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
Db        445 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 504

Qy        176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTA 235
Db        505 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 564

Qy        236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
Db        565 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 621

Qy        296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
Db        622 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 681

Qy        356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
Db        682 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 741

Qy        416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
Db        742 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 801

Qy        476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
Db        802 GGGAAGTATTTAAGACCTACAATGTCGCCGTGGACTACGTTACAGTAGCACTCCTAATCT 861

Qy        536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
Db        862 GGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 921

Qy        596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
Db        922 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 981

Qy        656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
Db        982 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 1041

Qy        716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
Db        1042 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1101

Qy        776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
Db        1102 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1158

Qy        836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
Db        1159 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1218

Qy        896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
Db        1219 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1278

Qy        955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
Db        1279 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1314

Qy        1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
Db        1315 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1370

Qy        1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
Db        1371 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1430

Qy        1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
Db        1431 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTA 1490

Qy        1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
Db        1491 GCCATACTGATCGGCCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTG 1550

Qy        1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
Db        1551 CCAGCCCTCCCCATCTCCATCACCTTCGGGCTCGTGTTCTACTT 1594


RESULT 43
AF149111
LOCUS       AF149111                1410 bp    mRNA    linear   ROD 01-JUN-2000
DEFINITION  Mus musculus presenilin-1 mRNA, complete cds.
ACCESSION   AF149111
VERSION     AF149111.1  GI:8131957
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1410)
  AUTHORS   Kumar,V.B., Vyas,K.C., Choudhary,V., Franko,M., Flood,J.F. and
            Morley,J.E.
  TITLE     Molecular Cloning and Tissue Distribution of Presenilin-1 in
            Senenscence Accelerated Mice (SAM P8) Mice
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1410)
  AUTHORS   Kumar,V.B., Vyas,K.C., Choudhary,V., Franko,M., Flood,J.F. and
            Morley,J.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-1999) Geriatric Research, St. Louis University/VA
            Medical Center, 11-G, GRECC, JB VA Medical Center, St. Louis, MO
            63125, USA
FEATURES             Location/Qualifiers
     source          1. .1410
                     /organism="Mus musculus"
                     /strain="SAM P8"
```

```
                     /db_xref="taxon:10090"
                     /tissue_type="hippocampus"
                     /dev_stage="4 months old"
     CDS             1. .1410
                     /codon_start=1
                     /product="presenilin-1"
                     /protein_id="AAF73153.1"
                     /db_xref="GI:8131958"
                     /translation="MTEIPAPLTYFQNAQMSEDSHSSSAIRSQNDSQERQQQHERQRL
                     DNPEPISNGRPQSNSRQVVEQDCMEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIK
                     SVSFYTRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVIMTILLVVLYKYRCY
                     KVIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYLTVALLIWNFGVVGMIAIHWKG
                     PLRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETA
                     QERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVPKNPKYNTQRARRDETQDSGSGND
                     DGGFSEEWEAQRDSHLGPHRSTPESRAAVQELSGSILTSEDPEERGVKLGLGDFIFYS
                     VLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFA
                     TDYLVQPFMDQLAFHQFYI"
BASE COUNT      359 a    362 c    341 g    348 t
ORIGIN

  Query Match          16.1%;  Score 240.8;  DB 10;  Length 1410;
  Best Local Similarity   53.6%;  Pred. No. 1.5e-44;
  Matches  635;  Conservative    0;  Mismatches  512;  Indels   37;  Gaps    5;

Qy        115 CAAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATT 174
              |||||         | |  |  |||||    |    |||||||  |||||     || || ||
Db        193 CAAGATTGTATGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATC 252

Qy        175 CATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATT 234
              || |||||  || ||| | || |||||||    | || ||     || || ||  |  |
Db        253 ATGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTC 312

Qy        235 ACGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACA 294
              |  || ||||  |  ||  | ||   ||   || | |     || ||       || ||
Db        313 AGCTTCTATACCCGGAAGGACGGTCAGC---TAATCTACACCCCATTCACAGAAGACACT 369

Qy        295 GACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTG 354
              || | | |   | | |   ||   ||  |    |||||   | |||||  |  | ||
Db        370 GAGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTC 429

Qy        355 GTCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATT 414
              | ||  | |||||  | || ||| | ||  | ||||||||||  ||  ||  ||| | ||
Db        430 ATTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATC 489

Qy        415 CATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTG 474
              ||||  |||||||||  |       | |  |  | || ||| | ||   |     ||  ||  |
Db        490 CATGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACCTA 549

Qy        475 CAAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTG 534
              |||||  | || |  | | ||||  |      |    ||   ||      |   | |
Db        550 GGGGAAGTATTTAAGACCTACAATGTCGCCGTGGACTACTTAACAGTAGCACTCCTAATC 609

Qy        535 GGTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAA 594
              | || | ||| ||  |||| |||||     || || |||||||| ||   | || |||||
Db        610 TGGAATTTTGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAG 669

Qy        595 CAGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCA 654
              |||   || || |||| ||   ||| || ||||| |||| ||||||||||||||| ||
Db        670 CAGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCC 729

Qy        655 GAATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTC 714
              ||||||||  |  ||| |  |  ||  ||| || || || |  ||| |||| || ||  |
Db        730 GAATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTA 789

Qy        715 ACACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATT 774
              || ||||| ||| |  |      |||| ||||| || ||||| ||||| ||| |  |
Db        790 TGTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTC 849

Qy        775 TTCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTT 834
              || || || || || ||||| ||   |  |  | |          | | | |
Db        850 TTTCCAGCTCTTATCTATTCCTCAACAATGGTGTGGTTGGTGAATATGGCT--------- 900

Qy        835 GAAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTT 894
```

```
                  ||  ||  |||   | || |         || |||| ||  | |
Db        901 --------GAAGGAGACCCAGAAGCCCAAAGGAGGGTACCCAAAAACCCCAAGTATAACA 952

Qy        895 CCTGGAGAGGCGAGTTGTTCATCTGAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
              |    |||| |  |     |    | | | ||      |    |  ||      |  ||
Db        953 CACAAAGAGCCAGGAGAGACGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCT 1012

Qy        955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
                |     |          | ||    ||||   |   |   | | ||  | |
Db        1013 TCAGTGAGGAGTGGGAGGCCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCACT---- 1068

Qy        1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
               |  |||  |  ||||     |       |    |    || | |    ||  |
Db        1069 ------CCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1122

Qy        1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
               ||   ||| || ||||| || |||||||| ||||| || ||||||||||||  ||||||
Db        1123 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1182

Qy        1135 CTCGGCAAGGCTTCATCGTACTT------TGACTGGAACACGACTATCGCTTGTTATGTG 1188
               | || ||||| ||| |   |          ||||||||||| || || || || | |||
Db        1183 GTTGGTAAGGCCTCAGCAACCGCAAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTA 1242

Qy        1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
               ||||| || ||||| || ||| | ||  |  ||||||||||| | |||||   |||  |
Db        1243 GCCATACTGATCGGCCTGTGCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTG 1302

Qy        1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
               || || ||  | || |||||    | ||| ||| | || |||||
Db        1303 CCAGCCCTCCCCATCTCCATCACTTTCGGCCTCGTGTTCTACTT 1346
```

```
RESULT 44
HSU40380
LOCUS       HSU40380                1484 bp    mRNA    linear   PRI 12-AUG-1996
DEFINITION  Human presenilin I-374 (AD3-212) mRNA, complete cds.
ACCESSION   U40380
VERSION     U40380.1  GI:1244639
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1484)
  AUTHORS   Sahara,N., Yahagi,Y., Takagi,H., Kondo,T., Okochi,M., Usami,M.,
            Shirasawa,T. and Mori,H.
  TITLE     Identification and characterization of presenilin I-467, I-463 and
            I-374
  JOURNAL   FEBS Lett. 381 (1-2), 7-11 (1996)
  MEDLINE   96193901
   PUBMED   8641442
REFERENCE   2  (bases 1 to 1484)
  AUTHORS   Shirasawa,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-1995) Takuji Shirasawa, Molecular Pathology,
            Tokyo Metropolitan Institute of Gerontology, 35-2 Sakae-cho,
            Itabashi-ku, Tokyo, 175, Japan
FEATURES             Location/Qualifiers
     source          1. .1484
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q24.3"
                     /clone="AD3-212"
     gene            1. .1484
                     /gene="AD3-212"
     CDS             1. .1125
                     /gene="AD3-212"
                     /codon_start=1
                     /product="presenilin I-374"
                     /protein_id="AAB05895.1"
                     /db_xref="GI:1244640"
                     /translation="MTELPAPLSYFQNAQMSEDNHLSNTNDNRERQEHNDRRSLGHPE
```

```
                     PLSNGRPQGNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKSVSFY
                     TRKDGQLIYTPFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHA
                     WLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQ
                     QAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNE
                     TLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNSKYNAERACLPPAAINLLSIAPMAPR
                     LFMPKGACRPTAQKGSHKTLLQRMMMAGSVRNGKPRGTVI"
     3'UTR           1126. .1484
                     /gene="AD3-212"
BASE COUNT      373 a    343 c    357 g    411 t
ORIGIN

  Query Match             15.9%;  Score 239;  DB 9;  Length 1484;
  Best Local Similarity   60.5%;  Pred. No. 3.8e-44;
  Matches  411;  Conservative    0;  Mismatches  265;  Indels    3;  Gaps    1;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178
              |||| ||  | |  |  || ||   |    ||||||| || ||    || || ||    |
Db        182 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 241

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238
              | ||||| || ||| | || |||||||   |||| ||   ||| || ||  |  | | .|
Db        242 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 301

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298
              ||||||  |  ||  |||   ||||  || | |      || ||          || || || |
Db        302 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 358

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
              | | |   |||  |   ||   ||| ||   ||||||  | ||||| |  | ||  | |
Db        359 CTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 418

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418
              || | |||||  |  || ||| | ||| | ||||||||||  ||  ||||||  | ||  ||||
Db        419 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 478

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478
              |||||||||  |      | |  |  | || ||| | ||  |    || ||  ||  |
Db        479 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 538

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538
              ||||  | ||||  |   | ||  ||      |   |||    |      | | |
Db        539 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 598

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598
              | | ||| ||  | |||||||| |  || || ||||||||||||  | || || || |||
Db        599 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 658

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658
              || || |||| ||  ||| || |||| |||| ||||||||||||||| || ||||
Db        659 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 718

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718
              ||||||  |||| |  |  ||  ||| || ||||| |  ||| | || || ||  |    |
Db        719 GGACTGCGTGGCTCATCTTGGCTGTGATTTCGGTATATGATTTAGTGGCTGTTTTGTGTC 778

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778
              | ||||| ||| |  |     |||| ||||| || ||||||||||| ||  |  |||| |
Db        779 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 838

Qy        779 CGGCGCTGATTTATTCGTC 797
              | || || ||||| || ||
Db        839 CAGCTCTCATTTACTCCTC 857


RESULT 45
AY043493
LOCUS       AY043493                1430 bp    mRNA    linear   VRT 02-SEP-2001
DEFINITION  Gallus gallus presenilin 2 mRNA, complete cds.
ACCESSION   AY043493
VERSION     AY043493.1  GI:15421098
KEYWORDS    .
SOURCE      Gallus gallus.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1430)
  AUTHORS   Korade Mirnics,Z., Keryanov,S., Lovelock,J. and Corey,S.J.
  TITLE     Cloning of chicken presenilins
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1430)
  AUTHORS   Korade Mirnics,Z., Keryanov,S., Lovelock,J. and Corey,S.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2001) Pediatrics, Children's Hospital of
            Pittsburgh, University of Pittsburgh, 3460 Fifth Avenue,
            Pittsburgh, PA 15213, USA
FEATURES             Location/Qualifiers
     source          1. .1430
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
     CDS             47. .1402
                     /note="transmembrane protein; putative gamma-secretase;
                     PS2"
                     /codon_start=1
                     /product="presenilin 2"
                     /protein_id="AAK95409.1"
                     /db_xref="GI:15421099"
                     /translation="MITFMNNSDSEDEPCNERTSLMSAESPPVPSYQDGLQASETREA
                     QTHRKRQTGSSRSPNNVADEDASDSDVRVRESALENEEEELTLKYGAKHVIMLFVPVT
                     LCMIVVVATIKSVRFYTEKNGQLIYTPFSEDTPSVGQRLLNSVLNTIIMISVIVVMTV
                     FLVVLYKYRCYKFIHGWLILSSFMLLFLFTYIYLGEVLKTYNVAMDYPTVILIIWNFG
                     AVGMIRIHWKGPLQLQQAYLIMISALMVLVFIKYLPEWSAWVILGAISIYDLIAVLCP
                     KGPLRMLXETAQERNQPIFPALIYSSAMIWTVGMAKPDTAAKGQSQQAWDAEDERENH
                     SSTSHSDSQILDTRSPAPSHPITLEEMEEEERGVKLGLGDFIFYSVLVGKAAATPSGD
                     WNTTLAXXVAILIGLCLTLLLLAVFKKALPALPISITFGLIFYFSTDNLVRTDPLEIS
                     V"
BASE COUNT      377 a    332 c    335 g    380 t      6 others
ORIGIN

  Query Match             15.9%;  Score 238.2;  DB 5;  Length 1430;
  Best Local Similarity   60.1%;  Pred. No. 5.8e-44;
  Matches  412;  Conservative    0;  Mismatches  270;  Indels    3;  Gaps    1;

Qy        124 GAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATCTATTT 183
              || ||||  | ||| ||   |   ||||||| ||||||   || || ||    || |||
Db        281 GAGAATGAAGAGGAGGAGCTGACTCTGAAATATGGAGCAAAGCATGTAATCATGCTTTTT 340

Qy        184 GTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGTTTTAT 243
              ||||| ||  | || || |||  | | || ||    || || ||  |  |    || |||
Db        341 GTGCCTGTCACCCTTTGTATGATTGTTGTCGTCGCCACCATAAAGTCAGTGCGCTTCTAT 400

Qy        244 AGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACAGTATC 303
              |   | || || |  ||||  | | |     |||||        ||| ||         |
Db        401 ACGGAGAAAAAACG---GGCAGCTGATCTACACCCCTTTCAGTGAGGACACCCCATCCGTG 457

Qy        304 GTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCGTTCTG 363
              |   ||      ||||  ||| |    ||  || || | ||| |  |||| |||||  | .
Db        458 GGCCAGCGTCTCTTGAACTCAGTGTTGAACACTATCATAATGATCAGCGTCATCGTTGTA 517

Qy        364 ATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATGGATGG 423
              |||||  ||  |  |  ||||  | |||||||||   |  ||||||  | ||  ||||| |||
Db        518 ATGACCGTGTTCTTAGTTGTGCTTTATAAATACCGCTGCTATAAGTTCATCCATGGTTGG 577

Qy        424 CTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAGAAGTT 483
              || ||  |   |  |||| |  || |||| |||||    || ||| |    |||||
Db        578 CTAATCCTGTCCTCTTTTATGTTGCTGTTCCTCTTCACCTACATATATCTCGGTGAAGTG 637

Qy        484 CTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTAACTAT 543
              ||||  |  | | |||||  |      |||  | ||     |  | ||  | |
Db        638 TTGAAGACATACAATGTGGCCATGGATTACCCCACAGTTATCTTAATCATCTGGAATTTT 697

Qy        544 GGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGTTCTAC 603
              |||| | | ||||||||  | || || ||||||||||| ||||  || || ||    ||
Db        698 GGAGCTGTTGGAATGATTCGCATCCACTGGAAAGGTCCCTTGCAGCTCCAGCAAGCATAT 757

Qy        604 CTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAATGGACT 663
```

```
Db        758 CTCATTATGATCAGTGCTCTGATGGTNCTGGTTTTCATTAAATATCTACCTGAGTGGTCA 817

Qy        664 GTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACACCAAAA 723

Db        818 GCATGGGTGATTCTAGGTGCAATCTCAATTTATGATCTGATTGCTGTTCTCTGTCCTAAG 877

Qy        724 GGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCCCGGCG 783

Db        878 GGGCCACTGAGAATGCTGGNAGAAACTGCACAGGAAAGAAACCAACCCATTTTTCCTGCA 937

Qy        784 CTGATTTATTCGTCTGGAGTCATCT 808

Db        938 TTAATATATTCCTCTGCTATGATAT 962
```

RESULT 46
AF416717

```
LOCUS       AF416717                1254 bp    mRNA    linear   PRI 21-OCT-2001
DEFINITION  Homo sapiens presenilin 1 (PSEN1) mRNA, complete cds.
ACCESSION   AF416717
VERSION     AF416717.1  GI:16303803
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1254)
  AUTHORS   Kang,L., Zhang,B., Zhou,Y., Peng,X., Yuan,J. and Qiang,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-SEP-2001) Department of Biochemistry, Institute of
            Basic Medical Science, Chinese Academy of Medical Scicenes, 5 Dong
            Dan San Tiao, Beijing 100005, P.R. China
FEATURES             Location/Qualifiers
     source          1. .1254
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .1254
                     /gene="PSEN1"
     CDS             1. .1137
                     /gene="PSEN1"
                     /codon_start=1
                     /product="presenilin 1"
                     /protein_id="AAL16811.1"
                     /db_xref="GI:16303804"
                     /translation="MTELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEHNDRRSL
                     GHPEPLSNGRPQGNSRQVVEQDEEEDEELTLKYGAKHVIMLFVPVTLCMVVVVATIKS
                     VSFYTRKDGQLIYTPFTEDTETVGQGALHSILNAAIMISVIVVMTILLVVLYKYRCYK
                     VIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMISIHWKGP
                     LRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQ
                     ERNETLFPALIYSSTMVWLVNMAEGDPEAQRRVSKNSKYNAERACLPPAAINLLSIAP
                     MAPRLFMPKGACRPTAQKGSHKTLLQRMMMAGSVRNGKPRGTVI"
BASE COUNT      316 a    287 c    322 g    329 t
ORIGIN
```

```
Query Match           15.8%;  Score 237.4;  DB 9;  Length 1254;
Best Local Similarity  60.4%;  Pred. No. 8.9e-44;
Matches  410;  Conservative    0;  Mismatches  266;  Indels    3;  Gaps    1;

Qy        119 AAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTCATC 178

Db        194 AAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAGCATGTGATCATGC 253

Qy        179 TATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTACGT 238

Db        254 TCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCGTGGCTACCATTAAGTCAGTCAGCT 313

Qy        239 TTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAGACA 298

Db        314 TTTATACCCGGAAGGATG---GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGA 370

Qy        299 GTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGGTCG 358
```

```
Db        371 CTGTGGGCCAGGGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTG 430

Qy        359 TTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTCATG 418

Db        431 TTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATG 490

Qy        419 GATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGCAAG 478

Db        491 CCTGGCTTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTCATTCATTTACTTGGGGG 550

Qy        479 AAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGGGTA 538

Db        551 AAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGA 610

Qy        539 ACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAACAGT 598

Db        611 ATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCCAGCAGG 670

Qy        599 TCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAGAAT 658

Db        671 CATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTTATCAAGTACCTCCCTGAAT 730

Qy        659 GGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCACAC 718

Db        731 GGACTGCGTGGCTCATCTTGGCTGTGATTTCGGTATATGATTTAGTGGCTGTTTTGTGTC 790

Qy        719 CAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTTTCC 778

Db        791 CGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAACGCTTTTTC 850

Qy        779 CGGCGCTGATTTATTCGTC 797

Db        851 CAGCTCTCATTTACTCCTC 869
```

RESULT 47
AR060025

```
LOCUS       AR060025                1929 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 3 from patent US 5840540.
ACCESSION   AR060025
VERSION     AR060025.1  GI:5986475
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1929)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Nucleic acids encoding presenilin II
  JOURNAL   Patent: US 5840540-A 3 24-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .1929
                     /organism="unknown"
BASE COUNT      489 a    494 c    490 g    447 t      9 others
ORIGIN
```

```
Query Match           15.6%;  Score 233.6;  DB 6;  Length 1929;
Best Local Similarity  53.6%;  Pred. No. 6.6e-43;
Matches  635;  Conservative    2;  Mismatches  505;  Indels   42;  Gaps    6;

Qy        116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175

Db        378 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 437

Qy        176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTA 235

Db        438 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 497

Qy        236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295

Db        498 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 554

Qy        296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
```

```
Db        555 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 614

Qy        356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415

Db        615 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 674

Qy        416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475

Db        675 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 734

Qy        476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535

Db        735 GGGAAGTATTTAAGACCTACAATGTCKCCGTGGACTACGTTACAGTAGCACTCCTAATCT 794

Qy        536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595

Db        795 GGAATTGGGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 854

Qy        596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655

Db        855 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 914

Qy        656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715

Db        915 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 974

Qy        716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775

Db        975 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1034

Qy        776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835

Db       1035 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1091

Qy        836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895

Db       1092 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1151

Qy        896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954

Db       1152 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1211

Qy        955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014

Db       1212 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1247

Qy       1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074

Db       1248 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1303

Qy       1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134

Db       1304 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1363

Qy       1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188

Db       1364 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTKGTA 1423

Qy       1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248

Db       1424 GCCATACTGATCGGCCTGTGCCTTANATTACTCCTGCTCGCCATTTACAAGAAAGGGTNG 1483

Qy       1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291

Db       1484 CCAGCCCNCCCCATCTCCATCACCTTCGGGTTCGTGTTCTNCTT 1527

RESULT 48
AR087180
LOCUS       AR087180               1929 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 3 from patent US 5986054.
ACCESSION   AR087180
VERSION     AR087180.1  GI:10013943
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1929)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Genetic sequences and proteins related to alzheimer's disease
  JOURNAL   Patent: US 5986054-A 3 16-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .1929
                     /organism="unknown"
BASE COUNT      489 a    494 c    490 g    447 t      9 others
ORIGIN

  Query Match           15.6%;  Score 233.6;  DB 6;  Length 1929;
  Best Local Similarity  53.6%;  Pred. No. 6.6e-43;
  Matches  635;  Conservative    2;  Mismatches  505;  Indels   42;  Gaps    6;

Qy        116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175

Db        378 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 437

Qy        176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTTACGATGAACACGATTA 235

Db        438 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 497

Qy        236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295

Db        498 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 554

Qy        296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355

Db        555 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 614

Qy        356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415

Db        615 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 674

Qy        416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475

Db        675 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 734

Qy        476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535

Db        735 GGGAAGTATTTAAGACCTACAATGTCKCCGTGGACTACGTTACAGTAGCACTCCTAATCT 794

Qy        536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595

Db        795 GGAATTGGGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 854

Qy        596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655

Db        855 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 914

Qy        656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715

Db        915 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 974

Qy        716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775

Db        975 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1034

Qy        776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835

Db       1035 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1091

Qy        836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895

Db       1092 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1151

Qy        896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954

Db       1152 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1211
```

```
Qy       955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
Db      1212 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1247

Qy      1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
Db      1248 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1303

Qy      1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
Db      1304 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1363

Qy      1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
Db      1364 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTKGTA 1423

Qy      1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
Db      1424 GCCATACTGATCGGCCTGTGCCTTANATTACTCCTGCTCGCCATTTACAAGAAAGGGTNG 1483

Qy      1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
Db      1484 CCAGCCCNCCCCATCTCCATCACCTTCGGGTTCGTGTTCTNCTT 1527


RESULT 49
AR134367
LOCUS       AR134367                1929 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 3 from patent US 6194153.
ACCESSION   AR134367
VERSION     AR134367.1  GI:14123272
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1929)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Methods for determining risk of developing alzheimer's disease by
            detecting mutations in the presenilin 1 (PS-1) gene
  JOURNAL   Patent: US 6194153-A 3 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .1929
                     /organism="unknown"
BASE COUNT      489 a    494 c    490 g    447 t      9 others
ORIGIN

  Query Match           15.6%;  Score 233.6;  DB 6;  Length 1929;
  Best Local Similarity   53.6%;  Pred. No. 6.6e-43;
  Matches  635;  Conservative    2;  Mismatches  505;  Indels   42;  Gaps    6;

Qy       116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
Db       378 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 437

Qy       176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTACGATGAACACGATTA 235
Db       438 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 497

Qy       236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
Db       498 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 554

Qy       296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
Db       555 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 614

Qy       356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
Db       615 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 674

Qy       416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
Db       675 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 734

Qy       476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
Db       735 GGGAAGTATTTAAGACCTACAATGTCKCCGTGGACTACGTTACAGTAGCACTCCTAATCT 794

Qy       536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
Db       795 GGAATTGGGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 854

Qy       596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
Db       855 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 914

Qy       656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
Db       915 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 974

Qy       716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
Db       975 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1034

Qy       776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
Db      1035 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1091

Qy       836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
Db      1092 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1151

Qy       896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
Db      1152 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1211

Qy       955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
Db      1212 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------------------- 1247

Qy      1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
Db      1248 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1303

Qy      1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
Db      1304 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1363

Qy      1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
Db      1364 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTKGTA 1423

Qy      1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
Db      1424 GCCATACTGATCGGCCTGTGCCTTANATTACTCCTGCTCGCCATTTACAAGAAAGGGTNG 1483

Qy      1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
Db      1484 CCAGCCCNCCCCATCTCCATCACCTTCGGGTTCGTGTTCTNCTT 1527


RESULT 50
AR144468
LOCUS       AR144468                1929 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 3 from patent US 6210919.
ACCESSION   AR144468
VERSION     AR144468.1  GI:15106335
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1929)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Genetic sequences and proteins related to alzheimer's disease
  JOURNAL   Patent: US 6210919-A 3 03-APR-2001;
```

```
FEATURES             Location/Qualifiers
     source          1. .1929
                     /organism="unknown"
BASE COUNT      489 a    494 c    490 g    447 t      9 others
ORIGIN
```

Query Match 15.6%; Score 233.6; DB 6; Length 1929;
Best Local Similarity 53.6%; Pred. No. 6.6e-43;
Matches 635; Conservative 2; Mismatches 505; Indels 42; Gaps 6;

```
Qy        116 AAGAAGACGAAAATGTTGTGGAAGAAGCGGAGCTGAAATACGGAGCATCTCACGTTATTC 175
Db        378 AACAAGATGAGGAGGAAGACGAAGAGCTGACATTGAAATATGGAGCCAAGCATGTCATCA 437

Qy        176 ATCTATTTGTGCCGGTGTCACTATGCATGGCTCTGGTTGTTTTTTACGATGAACACGATTA 235
Db        438 TGCTCTTTGTCCCCGTGACCCTCTGCATGGTCGTCGTCGTGGCCACCATCAAATCAGTCA 497

Qy        236 CGTTTTATAGTCAAAACAATGGAAGGCATTTACTATCACATCCTTTTGTCCGGGAAACAG 295
Db        498 GCTTCTATACCCGGAAGGACG---GTCAGCTAATCTACACCCCATTCACAGAAGACACTG 554

Qy        296 ACAGTATCGTTGAGAAGGGATTGATGTCACTTGGAAATGCTCTCGTCATGTTGTGCGTGG 355
Db        555 AGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATGCGGCCATCATGATCAGTGTCA 614

Qy        356 TCGTTCTGATGACAGTTCTGCTGATTGTTTTCTATAAATACAAGTTTTATAAGCTTATTC 415
Db        615 TTGTCATTATGACCATCCTCCTGGTGGTCCTGTATAAATACAGGTGCTACAAGGTCATCC 674

Qy        416 ATGGATGGCTTATTGTCAGCAGTTTTCTTCTTCTTTTCCTATTCACTACAATCTATGTGC 475
Db        675 ACGCCTGGCTTATTATTTCATCTCTGTTGTTGCTGTTCTTTTTTTCGTTCATTTACTTAG 734

Qy        476 AAGAAGTTCTGAAAAGTTTCGATGTGTCTCCCAGCGCACTATTGGTTTTGTTTGGACTGG 535
Db        735 GGGAAGTATTTAAGACCTACAATGTCKCCGTGGACTACGTTACAGTAGCACTCCTAATCT 794

Qy        536 GTAACTATGGAGTTCTCGGAATGATGTGTATACATTGGAAAGGTCCATTGCGTCTGCAAC 595
Db        795 GGAATTGGGGTGTGGTCGGGATGATTGCCATCCACTGGAAAGGCCCCCTTCGACTGCAGC 854

Qy        596 AGTTCTACCTTATTACAATGTCTGCACTAATGGCTCTGGTCTTTATCAAGTACCTACCAG 655
Db        855 AGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGGTATTTATCAAGTACCTCCCCG 914

Qy        656 AATGGACTGTGTGGTTTGTGCTGTTTGTTATCTCGGTTTGGGATCTGGTTGCCGTGCTCA 715
Db        915 AATGGACCGCATGGCTCATCTTGGCTGTGATTTCAGTATATGATTTGGTGGCTGTTTTAT 974

Qy        716 CACCAAAAGGACCATTGAGATATTTGGTGGAAACTGCACAGGAGAGAAACGAGCCAATTT 775
Db        975 GTCCCAAAGGCCCACTTCGTATGCTGGTTGAAACAGCTCAGGAAAGAAATGAGACTCTCT 1034

Qy        776 TCCCGGCGCTGATTTATTCGTCTGGAGTCATCTATCCCTACGTTCTTGTTACTGCAGTTG 835
Db        1035 TTCCAGCTCTTATCTATTCCTCAACAATGGTGT---GGTTGGTGAATATGGCTGAAGGAG 1091

Qy        836 AAAACACGACAGACCCCCGTGAACCGACGTCGTCAGACTCAAATACTTCTACAGCTTTTC 895
Db        1092 ACCCAGAAGCCCAAAGGAGGGTACCCAAGAACCCCAAGTATAACACACAAAGAGCGGAGA 1151

Qy        896 CTGGAGAGGCGAGTTGTTCATCT-GAAACGCCAAAACGGCCAAAAGTGAAACGAATTCCT 954
Db        1152 GAGAGACACAGGACAGTGGTTCTGGGAACGATGATGGTGGCTTCAGTGAGGAGTGGGAGG 1211

Qy        955 CAAAAAGTGCAAATCGAATCGAATACTACAGCTTCAACGACACAAAACTCTGGAGTAAGG 1014
Db        1212 CCCAAAGAGACAGTCACCTGGGGCCTCATCGCTCCA---------- 1247

Qy        1015 GTGGAACGGGAGCTAGCTGCTGAGAGACCAACTGTACAAGACGCCAATTTTCACAGGCAC 1074
Db        1248 ----CTCCCGAGTCAAGAGCTGCTGTCCAGGAACTTTCTGGGAGCATTCTAACGAGTGAA 1303

Qy        1075 GAAGAGGAAGAGAGAGGTGTGAAACTTGGTCTGGGCGACTTCATTTTCTACTCTGTTCTC 1134
Db        1304 GACCCGGAGGAAAGAGGAGTAAAACTTGGACTGGGAGATTTCATTTTCTACAGTGTTCTG 1363

Qy        1135 CTCGGCAAGGCTTCATCGTACT------TTGACTGGAACACGACTATCGCTTGTTATGTG 1188
Db        1364 GTTGGTAAGGCCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTKGTA 1423

Qy        1189 GCCATTCTTATCGGTCTCTGCTTCACTCTTGTCCTGCTCGCCGTCTTCAAACGAGCACTC 1248
Db        1424 GCCATACTGATCGGCCTGTGCCTTANATTACTCCTGCTCGCCATTTACAAGAAAGGGTNG 1483

Qy        1249 CCGGCTCT-GCAATTTCCATTTTCTCCGGACTCATTTTTTACTT 1291
Db        1484 CCAGCCCNCCCCATCTCCATCACCTTCGGGTTCGTGTTCTNCTT 1527

Search completed: July 15, 2003, 08:17:48
Job time : 4016 secs
```